TEXTBOOK

Wolfgang Cassing

Theoretical Physics Compact I

Classical Mechanics

Springer

Wolfgang Cassing

Theoretical Physics Compact I
Classical Mechanics

Springer

Wolfgang Cassing
University of Giessen, Gießen, Hessen, Germany

ISBN 978-3-031-95511-2 e-ISBN 978-3-031-95512-9
https://doi.org/10.1007/978-3-031-95512-9

This Springer imprint is published by the registered company Springer Nature Switzerland AG

The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Dedicated to Prof. Dr. Achim Weiguny

Preface

This book provides a textbook on classical mechanics and is in particular suited for bachelor students in their first year of studies in theoretical physics. The mathematical requirements include a knowledge of differentiation and integration and mathematical proofs are kept as simple as possible, however, still kept stringent. Elements of linear algebra are explained in detail in the text if needed in the context of coordinate transformations, rotations, Galilei or Lorentz transformations.

After defining the physical quantities of interest in the kinematics of mass points in inertial systems, the transformations between different inertial systems are derived. After these preparatory chapters, the Newtonian dynamics is formulated and examples for the solution of the equations of motion are presented. Furthermore, the tight connection between Galilei invariance and the conservation laws of momentum and angular momentum are pointed out. In case of conservative forces a potential energy can be formulated that—together with the kinetic energy of mass points—gives the energy of the system. The conservation of the total energy for a closed system follows in a straight forward manner. Applications of Newtonian mechanics for $1/r^2$-forces lead to Kepler's laws for the motion of planets and a gravitational field for a static mass distribution can be defined. Another important application is the harmonic oscillator being damped or driven by an external periodic force.

Since Maxwell's equations for electrodynamics are not Galilei invariant a new transformation law is derived (Lorentz transformation), which keeps the velocity of light c invariant in all inertial systems moving with relative velocity $v < c$. Some consequences are pointed out such as Lorentz contraction, time dilation, simultaneity, or causality of events. Mathematical aspects of Lorentz transformations are pointed out and the relativistic dynamics for mass points are derived accordingly. It is, furthermore, shown that the relativistic equations of motion merge with Newtonian dynamics for small velocities $v \ll c$.

The formal structure of mechanics is addressed in the second part of this book that aims at an algebraic formulation of the dynamics, which is independent on the particular choice of coordinates of an observer. After introducing generalized coordinates, that account for constraints on the system of particles and avoid the introduction of coercive forces, we introduce the Lagrange function and a variational principle to derive the Lagrange equations of motion. A Legendre transformation of the Lagrange function to the Hamilton function will lead to a description of the dynamics in phase-space variables, i.e. generalized coordinates and momenta. The Lagrange equations of motion turn to Hamilton's equations of motion, which can be expressed by Poisson brackets for the time evolution of an observable. The latter are shown to be invariant with respect to point transformations and extended canonical transformations of the phase-space variables such that a formal formulation of the classical mechanics is achieved that paves the way for the formulation of quantum mechanics, continuum mechanics, and statistical mechanics.

In the appendices the relativistic Lagrange and Hamilton functions for characteristic problems are given as well as numerical algorithms for differentiation and integration. Furthermore, algorithms of different order for the solution of differential equations are presented.

Acknowledgements This book results from the collaboration with many students and collaborators throughout about 35 years of common teaching and research. It follows the drafts of my teacher Prof. Dr. Achim Weiguny to whom this volume is dedicated. Special thanks go to my daughter Marie for preparing some of the figures and helpful comments on notations and presentations.

Wolfgang Cassing
Gießen, Germany
October 2024

About This Book

This book provides a textbook on classical mechanics and is in particular suited for bachelor students in their first year of studies in theoretical physics. The mathematical requirements include a knowledge of differentiation and integration; mathematical proofs are kept as simple as possible, however, still kept stringent.

The Newtonian dynamics are developed for systems of point masses and solved for a couple of characteristic examples. The extension to relativistic dynamics is outlined and the Lorentz transformation is derived in a simple case. Some consequences are pointed out such as Lorentz contraction, time dilation, simultaneity or causality of events.

The formal structure of mechanics is addressed in the second part of this book that aims at an algebraic formulation of the dynamics. The Lagrange and Hamilton functions are introduced and a variational principle is formulated, that leads to the Lagrange or Hamilton equations of motion. The latter are rewritten in terms of Poisson brackets in phase-space variables such that a formal formulation of the classical mechanics is achieved that paves the way for quantum mechanics, continuum mechanics and statistical mechanics.

The author is a retired Professor of Theoretical Physics at the university of Giesen and has shared the responsibility for the introduction of Bachelor and Master courses in Physics since 2005. His expertise is the phase-space dynamics of classical and quantum many-body systems, which in part is published in a book on transport theories. Moreover, he has written a series of textbooks in Theoretical Physics.

Contents

List of Figures

Fig. A.2 Illustration of an oscillating string in the continuum limit

W. Cassing, *Theoretical Physics Compact I*
https://doi.org/10.1007/978-3-031-95512-9_1

1. Overview

Wolfgang Cassing[1] ✉
(1) University of Giessen, Gießen, Hessen, Germany

1.1 Introduction

The description of phenomena in our daily world is a subtle problem since everybody has his personal point of view and different observers of the same phenomenon will provide different descriptions sometimes guided by personal preferences. Even the description of stationary objects depends on the position of the observer, the viewing angle and the relative motion of the observer who might be sitting in a car or train or be located in a rotating system. Furthermore, not all phenomena in our daily life are subject of a physical description and physical 'objects' need to be properly defined. Another mandatory requirement is that observations from observes in different systems must follow some transformation rules such that they can clearly specify if their observations are different or identical. A mathematical description is required to clearly define 'identical' results.

In this book we will start with the most simple systems, i.e. the motion of mass points in space and time and their trajectories under the influence of forces. The mathematical tools will be differentiation and integration in cartesian coordinate systems of a three-dimensional real vector space, which will be used to uniquely define physical quantities like inertial systems, velocity, acceleration, force, momentum, angular momentum or energy. A brief introduction to Euclidean vector spaces will be given and linear transformations (like rotations) be described by suitable 3×3 matrices. This will allow for a rigid formulation of kinematics in case of circular motion.

Once the physical quantities are defined it remains to clarify the conditions, that observers in different inertial systems—moving with some constant velocity $\overrightarrow{v_0}$ relative to each other—can find their observations to be identical. This leads us to the Galilean relativity principle and the Galilei group of transformations. Of special interest are rotating and center of mass systems which will be discussed in detail. After this preparatory work we will be able to define forces and derive Newton's equations of motion; their solution will provide the trajectory of a mass point in space and time. Examples for characteristic problems will be given and the explicit solutions derived in detail. It will turn out that instead of velocities or angular velocities it is more convenient to introduce momenta and angular momenta of particles since for closed systems—without external forces—the total momentum is a constant of motion. This also holds for the angular momentum if no external torque acts on the system. Next we will consider the connection between the work done by a force on a particle along its trajectory and the actual kinetic energy. In case of conservative forces we can introduce a potential energy $U(\overrightarrow{r})$ that allows to compute the actual force by its negative gradient. Then the energy of the system can be defined by the sum of kinetic and potential energy and—for closed systems—is found to be a conserved quantity, too.

We will continue with applications of Newtonian mechanics for central forces, where the potential *U* only depends on the magnitude of the relative distance $|\overrightarrow{r_1} - \overrightarrow{r_2}|$ between two mass points. In this case the conservation of momentum, angular momentum and energy holds which drastically reduces the number of free degrees of freedom. An important case are $1/r^2$-forces, which holds for Coulomb and gravitational forces; we will classify the trajectories according to their energy and derive Kepler's laws for the motion of planets. In extension the law of gravity is derived and gravity fields are introduced for static mass distributions. In addition the dynamics of a linear oscillator is discussed—another important physical system—and the solutions are computed from the equations of motion also in case of additional frictional forces. The case of a damped oscillator, that is driven by an external periodic force, will lead to the formation of resonances that are analysed in some detail. In addition the problem of coupled harmonic oscillations is addressed that is characteristic for the vibrational modes in crystals.

So far we have introduced classical Newton mechanics which, however, has different transformation properties than Maxwell's equations for electrodynamics. This incompatibility has been solved in Einstein's special theory of relativity. We thus have to replace the Galilei transformation between inertial systems by the Lorentz transformation that keeps the velocity of light c invariant in all inertial systems. We will derive the Lorentz transformation explicitly (in a simple case) and discuss its implications: Lorentz contraction, time dilation, simultaneity in moving systems as well as causality and the limiting velocity of signals. Some mathematical aspects of the Lorentz group of transformations will be discussed and Lorentz scalars, four-vectors and Lorentz tensors are identified as well as corresponding physical quantities like four-current densities. We close the discussion of relativistic dynamics by introducing the energy-momentum four-vector, which is conserved in all four components for closed systems, and discuss scattering problems. As an example the important problem of Compton scattering of a photon on a charge q is computed explicitly. The derivation of the Lorentz transformation of the force finalizes this chapter.

The equations of motion can be written in different ways—depending on the choice of coordinates—and in principle all independent choices have equal rights. However, some choices facilitate the solutions of the equations of motion and others might cause severe problems. It is thus of general interest to find 'optimal' coordinates for the description, which is also of practical help if the system is subject to constraints, that require the introduction of coercive forces which often are difficult to define. It is thus meaningful to define generalized coordinates that fulfill the constraints and also reduce the complexity of the problem by reducing the number of (linear independent) degrees of freedom. The equations of motion in generalized coordinates are derived from Newton's equations of motion. It is found that these equations can also be generated by a variational principle, which specifies a Lagrange function L, that is given by the difference between the kinetic and potential energy in case of conservative forces. An important consequence is that the Lagrange equations of motion can also be applied to other areas of physics. Generalized momenta are defined by the derivative of the Lagrange function with respect to the generalized velocity. Accordingly, if the Lagrange function does not depend on a specific coordinate, e.g. the azimuthal angle φ, the corresponding generalized momentum (here angular momentum) is a constant of motion. This suggests to transform the formulation to phase-space variables given by coordinates and their associated momenta, which is carried out by a Legendre transformation defining the Hamilton function H. In case of conservative forces the latter just gives the energy of the system in phase-space variables. The variational principle thus can be reformulated in terms of Hamilton's (equivalent) variational principle which gives the canonical equations of motion. The latter are illustrated for a couple of examples. Furthermore, it is shown that for a closed system the translational invariance leads to the conservation of the total momentum, the rotational invariance to the conservation of total angular momentum, and the invariance with respect to time translations to the conservation of the total energy.

Applications of the Lagrange formalism will be given for the motion of rigid bodies, which leads to the definition of an inertial tensor. The eigenvectors and eigenvalues of this tensor define the main axes of inertia and main moments of inertia, respectively. From the Lagrange function for the rigid body we will derive Euler's equation of motion, which are studied for the case of a symmetric heavy gyroscope.

Although the Lagrange formalism is a convenient method to tackle complex problems it is of advantage to formulate the dynamics in phase-space variables, i.e. in generalized coordinates and generalized momenta. In this case the time evolution of an observable, that not explicitly depends on time, is given by Poisson brackets which are determined by the derivative of the observable and the Hamiltonian with respect to the phase-space variables. The elementary Poisson bracket between generalized coordinates and generalized momenta turns out to be unity for associated pairs and their time evolution is given by the Poisson bracket with the Hamilton function, i.e. by the canonical equations of motion. The Poisson brackets thus allow for an algebraic formulation of the dynamics. However, the choice of generalized coordinates is not unique and invertible transformations between the coordinates are allowed. But not all transformations are meaningful since some transformations may lead to equations of motion that are no longer canonical. Allowed transformations then will be given by point transformations and extended canonical transformations that keep the equations of motion canonical invariant. Furthermore, the elementary Poisson brackets will be shown to be invariant with respect to canonical transformations such that a formulation of classical mechanics is achieved that is independent on the choice of the generalized coordinates. This will pave the way to quantum mechanics, where the Poisson brackets will be replaced by commutators of operators in an abstract Hilbert space. This also leads to a rigid formulation of statistical mechanics, where the physical system—in equilibrium—is described by ensembles with properties that are defined by expectation values of conserved observables.

In the appendices some useful extensions are presented: the Lagrange and Hamilton functions for relativistic systems as well as for continuum mechanics. We close by providing numerical algorithms for differentiation and integration as well as for the numerical solution of a set of differential equations.

1.2 Newton's Axioms

The starting points for classical non-relativistic mechanics[1] are **Newton's axioms** for the motion of a mass point (of mass m) under the influence of a force $\overrightarrow{F}$. By a mass point we understand a rigid body that possesses no **internal** degrees of freedom and can only perform translations (displacements) and rotations (turns).

The Newtonian axioms are explicitly stated as follows:

- **1st axiom**

  In an **inertial system** a **free** particle moves **colinear and uniform**.

- **2nd axiom**

  The state of motion of a particle of mass m changes under the influence of a **force** $\overrightarrow{F}$ according to

$$m \frac{d^2}{dt^2} \overrightarrow{r} = \overrightarrow{F}. \tag{1.1}$$

- **3rd axiom**

For the interaction between 2 mass points the **principle of action and reaction** applies, i.e.

$$\overrightarrow{F}_{12} = -\overrightarrow{F}_{21} \quad , \tag{1.2}$$

when $\overrightarrow{F}_{12}$ is the force exerted by particle 1 on particle 2.

- **4th axiom**

If two forces $\overrightarrow{F}_a$ and $\overrightarrow{F}_b$ act on a mass point, then the resulting force $\overrightarrow{F} = \overrightarrow{F}_a + \overrightarrow{F}_b$ has to be inserted into the equation of motion (**superposition principle** of forces).

The terms **free** particle, **inertial system**, and **force** require mathematical precision. A physical observation will always be meaningful when the statements made are independent of the observer, i.e. measurements in different reference systems can be compared and confirmed as identical. Mathematical tools—for the comparability of measurements—are provided in mechanics by the vector calculus and the theory of differential equations. Initially, however, it is expedient to introduce a series of simple concepts (also corresponding to natural intuition).

Footnotes

1 By **non-relativistic** we refer to all physical systems that move with velocities $v \ll c$, where $c \approx 300,000$ km/s denotes the velocity of light.

W. Cassing, *Theoretical Physics Compact I*
https://doi.org/10.1007/978-3-031-95512-9_2

2. Kinematics

Wolfgang Cassing[1]
(1) University of Giessen, Gießen, Hessen, Germany

In this chapter we will describe the motion of mass points in space and time and their trajectories by vectors $\vec{r}(t)$, $\vec{v}(t)$ and $\vec{a}(t)$. The mathematical tools are differentiation in cartesian or polar coordinate systems of a three-dimensional real vector space, which is used to uniquely define physical quantities like inertial systems, velocity, acceleration, angular velocity or angular acceleration. A brief introduction to Euclidean vector spaces will been given and linear transformations (like rotations) be described by suitable 3×3 matrices. This allows for a rigid formulation of kinematics also in case of circular motion.

2.1 Basic Terms

2.1.1 Straight-Line Motion

To describe the straight-line motion we select a cartesian coordinate system such that the mass point e.g. is moving (in a single dimension) along the x-axis (see Fig. 2.1).

The sequence of motions is determined by the position x of the mass point at time t $(x = x(t))$. The trajectory $x(t)$ in this case is completely determined.

We define the **average velocity** by

$$v_m = \frac{x(t')-x(t)}{t'-t} = \frac{\Delta x}{\Delta t}, \tag{2.1}$$

where Δx is the displacement during the time interval Δt.

If $x(t)$ is differentiable with respect to t we define the **velocity** $v(t)$ by

$$v(t) = \lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} = \frac{dx}{dt}. \tag{2.2}$$

If the velocity v remains constant during the entire motion, i.e. v is independent of t, we call the motion **uniform**.

The **average acceleration**, furthermore, is defined by

$$a_m(t) = \frac{v(t')-v(t)}{t'-t} = \frac{\Delta v}{\Delta t}. \tag{2.3}$$

If $x(t)$ is at least twice differentiable with respect to t the **acceleration** then is given by

$$a(t) = \lim_{\Delta t \to 0} \frac{\Delta v}{\Delta t} = \frac{dv}{dt} = \frac{d^2x}{dt^2}. \tag{2.4}$$

If $a \neq 0$ is independent of time t we call the motion **uniformly accelerated**.

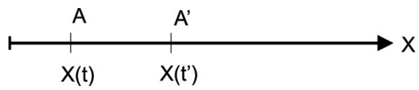

Fig. 2.1 Illustration for straight-line motion in a single dimension (along the *x*-axis)

2.1.2 Curved Motion

We describe the position of a particle on its trajectory (in 3 spatial dimensions) by its coordinates *x*, *y*, *z* in a **cartesian** coordinate system. We define a **coordinate vector**

$$\overrightarrow{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \tag{2.5}$$

which points from the coordinate origin to the position $\overrightarrow{P}$ of the particle. The sequence of motions is then determined by the functions

$$x = x(t), \;\; y = y(t), \;\; z = z(t) \tag{2.6}$$

or in vector notation

$$\overrightarrow{r} = \overrightarrow{r}(t) \tag{2.7}$$

The **average velocity** then is given by:

$$\overrightarrow{v}_m = \frac{\overrightarrow{r}(t') - \overrightarrow{r}(t)}{t' - t} = \frac{\Delta\overrightarrow{r}}{\Delta t} = \begin{pmatrix} \frac{\Delta x}{\Delta t} \\ \frac{\Delta y}{\Delta t} \\ \frac{\Delta z}{\Delta t} \end{pmatrix}. \tag{2.8}$$

It is represented by a vector in the direction of the **displacement vector** $\Delta\overrightarrow{r}$.

If the functions $x(t), \;\; y(t), \;\; z(t)$ are differentiable with respect to time *t*, the **velocity** is defined by:

$$\overrightarrow{v} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \lim_{\Delta t \to 0} \frac{\Delta\overrightarrow{r}}{\Delta t} = \frac{d\overrightarrow{r}}{dt}. \tag{2.9}$$

The velocity $\overrightarrow{v}$ is represented by a vector in direction of the tangent to the trajectory at point $\overrightarrow{P}$.

The **length** of the position vector $\overrightarrow{r}$ is given by:

$$|\overrightarrow{r}| = r = \sqrt{x^2 + y^2 + z^2}, \tag{2.10}$$

the **magnitude of the velocity** by:

$$|\overrightarrow{v}| = v = \sqrt{v_x^2 + v_y^2 + v_z^2}. \tag{2.11}$$

If the functions $v_x(t), v_y(t), v_z(t)$ are all differentiable with respect to *t*, the **acceleration** $\overrightarrow{a}$ becomes

$$\vec{a} = \lim_{\Delta t \to 0} \frac{\Delta \vec{v}}{\Delta t} = \frac{d\vec{v}}{dt} = \frac{d^2\vec{r}}{dt^2}. \tag{2.12}$$

Note: More than 2nd derivatives of the trajectory $\vec{r}(t)$ with respect to time t are not needed because in Newton's equations of motion at most 2nd derivatives appear.

2.1.3 Curvature of Trajectories

The **velocity** $\vec{v}$ is a vector in the direction of the tangent to the trajectory. We can therefore also write

$$\vec{v}(t) = v(t)\vec{e_T}(t); \qquad \vec{e_T}(t) = \frac{\vec{v}(t)}{|\vec{v}(t)|}, \tag{2.13}$$

with $\vec{e_T}$ as **unit vector** in the direction of the respective tangent to the trajectory.

The **acceleration** $\vec{a}$ (according to the product rule of differentiation) then reads:

$$\vec{a} = \frac{d}{dt}\left(v(t)\vec{e_T}(t)\right) = \underbrace{\frac{dv}{dt}\vec{e_T}}_{1.} + \underbrace{v\frac{d\vec{e_T}}{dt}}_{2.} \tag{2.14}$$

The acceleration thus can be separated in two components:

1. the **tangential component** ($\sim \vec{e_T}(t)$):

$$\vec{a_T} = \frac{dv}{dt}\vec{e_T} \tag{2.15}$$

2. and the **normal component** which is perpendicular to $\vec{e_T}$ and given by:

$$\vec{a_N} = v\frac{d\vec{e_T}}{dt}. \tag{2.16}$$

A useful **component representation** of $\vec{e_T}$ is:

$$\vec{e_T}(t) = \begin{pmatrix} \cos\varphi(t) \\ \sin\varphi(t) \\ 0 \end{pmatrix}, \tag{2.17}$$

which gives

$$\frac{d\vec{e_T}}{dt} = \begin{pmatrix} -\dot{\varphi}\sin\varphi \\ \dot{\varphi}\cos\varphi \\ 0 \end{pmatrix} = \dot{\varphi}\begin{pmatrix} \cos\left(\varphi + \frac{\pi}{2}\right) \\ \sin\left(\varphi + \frac{\pi}{2}\right) \\ 0 \end{pmatrix} = \dot{\varphi}\vec{e_N}. \tag{2.18}$$

With the abbreviation $d\varphi/dt = \dot{\varphi}$ we get

$$\vec{a} = \vec{a_T} + \vec{a_N} \tag{2.19}$$

with

$$\vec{a}_N = v\dot{\varphi}\vec{e}_N. \tag{2.20}$$

The magnitude of $\dot{\varphi}$ is closely related to the curvature of the trajectory. The **arc length** $s = s(t)$ depends on the magnitude of the velocity via

$$\frac{ds}{dt} = v. \tag{2.21}$$

Using the chain rule we obtain

$$\boxed{\dot{\varphi} = \frac{d\varphi}{dt} = \frac{d\varphi}{ds}\frac{ds}{dt} = \frac{d\varphi}{ds}v.} \tag{2.22}$$

The quantity introduced in this way can be calculated as $d\varphi/ds$ and interpreted geometrically: The intersection of the path normals of neighboring points A, A' in the limit $\Delta t \to 0$ is called **center of curvature**.

For the corresponding **curvature radius** $\varrho = \varrho(t)$ we get:

$$\frac{1}{\varrho} = \lim_{\Delta t \to 0} \frac{\Delta\varphi}{\Delta s} = \frac{d\varphi}{ds} \tag{2.23}$$

$$\to \vec{a}_N = \frac{v^2}{\rho}\vec{e}_N. \tag{2.24}$$

Special cases:

1. colinear motion:

   $$\varrho \to \infty, \qquad \text{i. e. } a_N \to 0 \tag{2.25}$$

2. circular motion:

   $$\varrho = R_{\text{circle}} = \text{const.} \tag{2.26}$$

After these rather clear definitions it is now important to clarify for which conditions 2 observers in different systems Σ and Σ' measure the **same** trajectories $\vec{r}(t), \vec{r}\,'(t)$ or denote them as **identical**. To this aim we first briefly recall basic elements of vector analysis.

2.2 Vectors

2.2.1 Definition

We define a vector $\vec{a}$ in $\mathbb{R}^3$[1] by a triple of real numbers $a_1,\ a_2,\ a_3$ (components) and write

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}. \tag{2.27}$$

We call two vectors $\vec{a}, \vec{b}$ **equal** if and only if:

(2.28)

$$a_1 = b_1 \qquad a_2 = b_2 \qquad a_3 = b_3.$$

2.2.2 Real Vector Spaces

In real vector spaces an **addition** $(+)$ of vectors is defined as well as a **multiplication** of vectors with real numbers.

The **addition** of 2 vectors $\vec{a}, \vec{b}$:

$$\vec{a} + \vec{b} = \vec{c} \tag{2.29}$$

is defined by

$$a_1 + b_1 = c_1 \qquad a_2 + b_2 = c_2 \qquad a_3 + b_3 = c_3. \tag{2.30}$$

The addition introduced in this way assigns exactly a single vector to every two vectors and has the following properties:

1. **commutativity**

$$\vec{a} + \vec{b} = \vec{b} + \vec{a} \tag{2.31}$$

2. **associativity**

$$(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c}) \tag{2.32}$$

3. **neutral element** There is a vector $\vec{0}$ with the property

$$\vec{a} + \vec{0} = \vec{a} \tag{2.33}$$

for any vector $\vec{a}$, i.e. the vector with the components (0, 0, 0).

4. **inverse element** For every vector $\vec{a}$ with components a_1, a_2, a_3 there is exactly one vector $(-\vec{a})$ such that

$$\vec{a} + (-\vec{a}) = \vec{0}, \tag{2.34}$$

i.e. the vector with the components $(-a_1, -a_2, -a_3)$.

Elements (here: vectors) with a connection (here: addition rule) with the properties 1 to 4 form a **commutative group**.

Furthermore, we define the

multiplication of vectors with real numbers α by:

$$\alpha\vec{a} = \begin{pmatrix} \alpha a_1 \\ \alpha a_2 \\ \alpha a_3 \end{pmatrix}. \tag{2.35}$$

The multiplication has the following properties:

1. **associativity**

$$(\alpha\beta)\vec{a} = \alpha(\beta\vec{a}) \qquad (2.36)$$

2. **distributivity**

$$(\alpha + \beta)\vec{a} = \alpha\vec{a} + \beta\vec{a} \qquad (2.37)$$

$$\alpha(\vec{a} + \vec{b}) = \alpha\vec{a} + \alpha\vec{b}$$

for arbitrary real numbers α, β.

3. **neutral element**:

$$1\vec{a} = \vec{a}. \qquad (2.38)$$

A commutative group, where elements are multiplied by real numbers and have the properties 1–3, defines a **real vector space**. The position vectors $\vec{r}$ and displacement vectors $\Delta\vec{r}$ form such a real vector space (in 3 dimensions $\mathbb{R}^3$).

2.2.3 Euclidean Vector Spaces

In Euclidean vector spaces the **length** of vectors can be defined as well as an **angle** between 2 vectors.

In 3-dimensional position space, the **length** (or **norm**) of a position vector is given by

$$r = |\vec{r}| = \sqrt{x^2 + y^2 + z^2} \geq 0 \qquad (2.39)$$

and the **angle** φ between arbitrary 2 position vectors is determined by

$$|\vec{r}_1 - \vec{r}_2|^2 = r_1^2 + r_2^2 - 2r_1r_2 \cos\varphi. \qquad (2.40)$$

These properties characterize an **Euclidean space**.

Mathematically, one gets an Euclidean vector space from a real vector space in the following way: One defines—between two vectors, $\vec{a}$ and $\vec{b}$—a **scalar product** (or **dot product**) $\vec{a} \cdot \vec{b}$ with the following properties:

1. $\vec{a} \cdot \vec{b}$ is a real number
2. $\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$ (commutative)
3. $(\alpha\vec{a}) \cdot \vec{b} = \alpha(\vec{a} \cdot \vec{b})$ (associative)

4. $\overrightarrow{a} \cdot (\overrightarrow{b} + \overrightarrow{c}) = \overrightarrow{a} \cdot \overrightarrow{b} + \overrightarrow{a} \cdot \overrightarrow{c}$ (distributive)

5. $\overrightarrow{a} \cdot \overrightarrow{a} = |\overrightarrow{a}|^2 \begin{cases} = 0 \text{ if } \overrightarrow{a} = \overrightarrow{0} \\ > 0 \text{ else} \end{cases}$

By

$$\cos\varphi = \frac{\overrightarrow{a} \cdot \overrightarrow{b}}{|\overrightarrow{a}||\overrightarrow{b}|} \tag{2.41}$$

we can introduce an angle φ, which turns out to be the intermediate angle of $\overrightarrow{a}$ and $\overrightarrow{b}$.

Using the scalar product we can also define the **orthogonality of vectors**:

2 vectors $\overrightarrow{a}, \overrightarrow{b}$ are called orthogonal to each other if:

$$\overrightarrow{a} \cdot \overrightarrow{b} = 0. \tag{2.42}$$

Geometrically the two vectors then are **perpendicular** to each other.

2.2.4 Basis and Dimension of Vector Spaces

In order to define a basis of a vector space we need the concept of **linear independence**:

vectors $\overrightarrow{a}_1, \ \overrightarrow{a}_2, \ldots, \ \overrightarrow{a}_i$ are called **linearly independent**, if in the vector addition

$$\alpha_1 \overrightarrow{a}_1 + \alpha_2 \overrightarrow{a}_2 + \ldots + \alpha_i \overrightarrow{a}_i = \overrightarrow{0} \tag{2.43}$$

for all real coefficients α_k $(k = 1, .., i)$ always follows

$$\alpha_1 = \alpha_2 = \cdots = \alpha_i = 0; \tag{2.44}$$

otherwise the vectors are denoted to be **linearly dependent**.

The unit vectors

$$\overrightarrow{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \overrightarrow{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \overrightarrow{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{2.45}$$

are linearly independent, because from

$$\alpha_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{2.46}$$

necessarily follows $\alpha_1 = \alpha_2 = \alpha_3 = 0$. Using the vectors (2.45) any vector $\overrightarrow{a}$ can be represented by a linear combination of the basis vectors:

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + a_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{2.47}$$

In short:

$$\overrightarrow{a} = a_1\overrightarrow{e_1} + a_2\overrightarrow{e_2} + a_3\overrightarrow{e_3} = \sum_{i=1}^{3} a_i\overrightarrow{e_i}. \tag{2.48}$$

The **basis of a vector space** is a set of linearly independent vectors covering the entire vector space such that **every** vector of the considered vector space **uniquely** can be written as a linear combination of the **basis vectors**. The number of basis vectors for a given vector space is fixed and defines the **dimension of the vector space**. The vectors (2.45) form a basis of the vector space of dimension 3.

Of particular practical importance (in physics) are vectors $\overrightarrow{e_i}$, which form an **orthonormal basis**.

They have the property

$$\overrightarrow{e_i} \cdot \overrightarrow{e_k} = \delta_{ik} \tag{2.49}$$

with the abbreviation:

$$\delta_{ik} = \begin{cases} 1 \text{ for } i = k \\ 0 \text{ for } i \neq k. \end{cases} \tag{2.50}$$

3 orthonormal vectors thus form a basis of a 3-dimensional vector space.

Note: When using an orthonormal basis, the dot product gets a particularly simple explicit form. Let a_i, b_i $(i = 1, 2, 3)$ be the components (also called "coordinates") of 2 vectors $\overrightarrow{a}, \overrightarrow{b}$ with respect to an orthonormal basis $\overrightarrow{e_k}$ $(k = 1, 2, 3)$

$$\overrightarrow{a} = \sum_{k=1}^{3} a_k\overrightarrow{e_k} \tag{2.51}$$

$$\overrightarrow{b} = \sum_{i=1}^{3} b_i\overrightarrow{e_i} \tag{2.52}$$

the scalar product becomes:

$$\overrightarrow{a} \cdot \overrightarrow{b} = \left(\sum_{k=1}^{3} a_k\overrightarrow{e_k}\right) \cdot \left(\sum_{i=1}^{3} b_i\overrightarrow{e_i}\right) = \sum_{k=1}^{3} \sum_{i=1} a_k b_i (\overrightarrow{e_k} \cdot \overrightarrow{e_i}) = \sum_{k=1}^{3} \sum_{i=1}^{3} a_k b_i \delta_{ki} = \sum_{i=1}^{3} a_i b_i. \tag{2.53}$$

Furthermore, the dot (or scalar) product of $\overrightarrow{a}$ and $\overrightarrow{e_k}$ gives the components a_k of $\overrightarrow{a}$ with respect to $\overrightarrow{e_k}$:

$$\overrightarrow{a} \cdot \overrightarrow{e_k} = \sum_{i=1}^{3} a_i (\overrightarrow{e_i} \cdot \overrightarrow{e_k}) = \sum_{i=1}^{3} a_i \delta_{ik} = a_k. \tag{2.54}$$

To illustrate these results, let us consider a position vector $\overrightarrow{r}$, given by its coordinates *x*, *y*, *z* in a cartesian coordinate system:

$$\overrightarrow{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x\overrightarrow{e_x} + y\overrightarrow{e_y} + z\overrightarrow{e_z}. \tag{2.55}$$

Here $\vec{e}_x, \vec{e}_y, \vec{e}_z$ are unit vectors in direction of the mutually orthogonal axes (**cartesian basis**),

$$\vec{e}_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \vec{e}_y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \vec{e}_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{2.56}$$

Then the length squared of $\vec{r}$ is given by

$$\vec{r} \cdot \vec{r} = |\vec{r}|^2 = x^2 + y^2 + z^2, \tag{2.57}$$

and the length by

$$|\vec{r}| = r = \sqrt{x^2 + y^2 + z^2}. \tag{2.58}$$

The dot product

$$\vec{r} \cdot \vec{e}_x = x \tag{2.59}$$

gives the length of the vector in *x*-direction by **orthogonal projection**. The same holds for the *y*- and *z*-direction when taking the scalar product with the respective basis vector $\vec{e}_y$ or $\vec{e}_z$.

2.3 Orthogonal Transformations

2.3.1 Vectors in Mathematics and Physics

While in mathematics vectors are simply elements of an (arbitrary) vector space, in physics a vector space is always understood as elements of Euclidean vector spaces!

When exposing two position vectors to a rotation in space or a reflection at the origin, the length and intermediate angle do not change!

2.3.2 Rotations

We now study the change of the components of a position vector $\vec{r}$ when the coordinate system rotates around the *z*-axis by an angle φ. We find:

$$\begin{aligned} x' &= x \cos\varphi + y \sin\varphi \\ y' &= -x \sin\varphi + y \cos\varphi \\ z' &= z. \end{aligned} \tag{2.60}$$

With the notation

$$x = x_1\ ,\ y = x_2\ ,\ z = x_3\ ;\ x' = x_1{}'\ ,\ y' = x_2{}'\ ,\ z' = x_3{}' \tag{2.61}$$

we can write in compact form

$$x_i{}' = \sum_{j=1}^{3} d_{ij} x_j \quad ; i = 1, 2, 3\ , \tag{2.62}$$

where the rotation **matrix** (d_{ij}) has the form:

$$(d_{ij}) = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.63}$$

Remark:

- For an arbitrary rotation the connection between the coordinates x_i and $x_j{}'$ is linear again, but the matrix (d_{ij}) has a more complicated form.

General properties of the matrix for a rotation:

Since during rotations the length of vectors and the angle between each of two vectors cannot change, the dot product must be invariant under rotations.

For 2 vectors $\overrightarrow{r_1}$, $\overrightarrow{r_2}$ with the components

$$\begin{array}{lcll} x_{1i} & , & x_{2i} & \text{in the system } XYZ \\ x'_{1j} & , & x'_{2j} & \text{in the system } X'Y'Z' \end{array} \tag{2.64}$$

must hold:

$$\overrightarrow{r'_1} \cdot \overrightarrow{r'_2} = \sum_{i=1}^{3} x_{1i}{}' x_{2i}{}' = \sum_{i=1}^{3} \left(\sum_{m=1}^{3} d_{im} x_{1\,m} \right) \left(\sum_{n=1}^{3} d_{in} x_{2n} \right) = \sum_{n=1}^{3} x_{1n} x_{2n} = \overrightarrow{r_1} \cdot \overrightarrow{r_2}. \tag{2.65}$$

It follows that

$$\sum_{i=1}^{3} d_{im} d_{in} = \sum_{i=1}^{3} d^T_{mi} d_{in} = \delta_{mn} \tag{2.66}$$

for the invariance of the scalar product during the transformation. Linear transformations with the property (2.66) are called **orthogonal transformations**.

2.3.3 Reflection at the Origin (Inversion)

We now consider the discrete transformation

$$x_i \;\to\; x_i{}' \;=\; -x_i. \tag{2.67}$$

The corresponding transformation matrix ($x'_i = \sum_k s_{ik} x_k$) has the form

$$(s_{ik}) \;=\; \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{2.68}$$

The difference between the orthogonal transformations presented here is that during rotations a right-handed system remains right-handed, while in an inversion it goes over to a left-handed system. This is expressed mathematically by the determinant of the transformation. The difference is that for rotations always holds:

$$\det\ (d_{ik}) = 1, \tag{2.69}$$

while in the case of reflection we have

$$\det\ (s_{ik}) = -1. \tag{2.70}$$

Remark: The reflection on a plane, e.g.

$$x'_1 = x_1 \; ; \; x'_2 = x_2 \; ; \; x'_3 = -x_3 \; , \tag{2.71}$$

can be done by combining the reflection at the origin and a rotation around the z-axis.

Appendum: Elementary determinants
We recall that the determinant of a square matrix a_{ik} for 2×2 matrices is defined as:

$$\det\,(a_{ik}) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}, \tag{2.72}$$

and for 3×3 matrices by:

$$\det\,(a_{ik}) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$

$$a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \tag{2.73}$$

The following rules are useful for practical calculations:

- **Rule 1**:

$$\det\,(A) = \det\,(a_{ik}) \; = \det\,(a_{ki}) = \det\,(A^T) \tag{2.74}$$

- **Rule 2**:

  If we swap in the matrix 2 rows (columns) the determinant changes sign.

  Conclusion:

  If 2 rows (columns) in a matrix are the same (or differ by a constant factor), then the determinant is zero.

- **Rule 3**:

  If we add a multiple of another row (column) to a row (column) the determinant does not change.

Note: The determinant of a matrix A (or linear transformation) is important for the existence of the **inverse** matrix A^{-1}. The latter only exists if det $A \neq 0$, i.e. det $(A^{-1}A)$ = det A det $A^{-1} = 1$.

2.3.4 Vectors and Scalars

We can now (in non-relativistic physics) define **vectors** as (ordered) triples of real numbers, for which

1. an addition and multiplication is defined in line with Sect. 2.2.2.
2. and which behave like position vectors $\overrightarrow{r_i}$ during rotations.

Note: The velocity $\overrightarrow{v}$ and acceleration $\overrightarrow{a}$ are vectors. The vectors (position vectors, velocity, acceleration), which change the sign by reflection, are called **polar** vectors. This also holds for momenta and forces (see following chapters).

The vectors, which do not change the sign by reflection, are called **axial** vectors. Examples for axial vectors are: angular momentum, torque (see following chapters).

2.3.5 Benefits of the Vector Calculation

Simplification of notation: Instead of specifying the components $x(t)$, $y(t)$, $z(t)$, one writes shorter: $\overrightarrow{r}(t)$.

Independence of the coordinate system: Statements in the form of vector equations hold regardless of the choice of the coordinate system.

2.4 Circular Motion

2.4.1 Angular Velocity

We consider the motion of a mass point on a circle with radius r. A useful parameter representation of the trajectory is given by:

$$\boxed{\overrightarrow{r}(t) = r\begin{pmatrix}\cos\,\varphi(t)\\ \sin\,\varphi(t)\\ 0\end{pmatrix} \qquad (2.75)}$$

with $r =$ const and the center of the circle as the origin.

The **velocity**

$$\boxed{\overrightarrow{v} = \frac{d\overrightarrow{r}}{dt} = r\dot{\varphi}\begin{pmatrix}-\,\sin\,\varphi(t)\\ \cos\,\varphi(t)\\ 0\end{pmatrix} = r\dot{\varphi}\overrightarrow{e_T} \qquad (2.76)}$$

has the **magnitude**

$$v = |v| = r\dot{\varphi} \qquad (2.77)$$

and is always directed perpendicular to $\overrightarrow{r}$, since

$$\overrightarrow{r}\cdot\overrightarrow{v} = r^2\dot{\varphi}(-\,\cos\,\varphi\,\sin\,\varphi + \sin\,\varphi\,\cos\,\varphi) = 0. \qquad (2.78)$$

The magnitude of the **angular velocity** ω is introduced via

$$\omega = \dot{\varphi} = \frac{v}{r}. \qquad (2.79)$$

If the position vector $\overrightarrow{r}$ of any mass point of the rotating body is not in the orbital plane of the mass point (see Fig. 2.2), (2.79) has to be replaced by:

$$v = r_0\dot{\varphi} = r\dot{\varphi}\,\sin\,\gamma = r\omega\,\sin\,\gamma. \qquad (2.80)$$

We can characterize any rigid rotation by the vector **angular velocity** $\overrightarrow{\omega}$, whose magnitude is determined by equation (2.80) and its direction is parallel to the axis of rotation in the sense of a **right-hand screw**

Fig. 2.3.

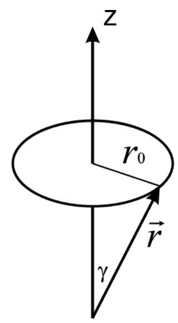

Fig. 2.2 Illustration for circular motion if the origin of the coordinate system is not in the plane of motion

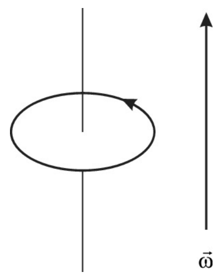

Fig. 2.3 Direction of the angular velocity $\vec{\omega}$ (right-handed)

The general connection of $\vec{r}, \vec{v}$ and $\vec{\omega}$ is described by the

2.4.2 Vector Product

The vector product of 2 vectors $\vec{a}, \vec{b}$ is defined as a vector $\vec{c}$, written as

$$\boxed{\vec{c} = \vec{a} \times \vec{b}, \qquad (2.81)}$$

its length

$$c = |\vec{c}| = ab \sin \gamma \qquad (2.82)$$

with γ defined by the angle between $\vec{a}$ and $\vec{b}$ and whose direction is perpendicular to $\vec{a}$ and $\vec{b}$, i.e.

$$\vec{a} \cdot \vec{c} = 0 \qquad \vec{b} \cdot \vec{c} = 0 \qquad (2.83)$$

in such a way that $\vec{a}, \vec{b}, \vec{c}$ give a right-handed system. The components of the vector $\vec{c}$ then are (as a function of the components of $\vec{a}$ and $\vec{b}$) given by:

$$\boxed{\vec{c} = \begin{pmatrix} a_y b_z - a_z b_y \\ a_z b_x - a_x b_z \\ a_x b_y - a_y b_x \end{pmatrix}. \qquad (2.84)}$$

Properties of the vector product :

1. Anticommutativity:

$$\vec{a} \times \vec{b} = -\vec{b} \times \vec{a} \qquad (2.85)$$

2. If $\vec{a}$ is parallel to $\vec{b}$, then

$$\vec{a} \times \vec{b} = \vec{0} \qquad (2.86)$$

3. Associative law: $(\alpha \in \mathbb{R})$

$$(\alpha\vec{a}) \times \vec{b} = \alpha(\vec{a} \times \vec{b}) \qquad (2.87)$$

4. Distributive law:

$$\vec{a} \times (\vec{b}_1 + \vec{b}_2) = \vec{a} \times \vec{b}_1 + \vec{a} \times \vec{b}_2 \qquad (2.88)$$

Geometric interpretation of $|\vec{a} \times \vec{b}|$: The area of the parallelogram formed by $\vec{a}$ and $\vec{b}$ (see Fig. 2.4) is given by:

$$A = |\vec{a} \times \vec{b}| = ab \sin\gamma \qquad \text{for } 0 \leq \gamma \leq \pi \qquad (2.89)$$

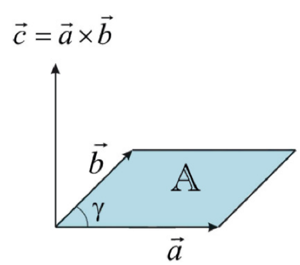

Fig. 2.4 Illustration of the parallelogram formed by vectors $\vec{a}$ and $\vec{b}$

Calculation rules:

For any vectors $\vec{a}, \vec{b}, \vec{c}$ the following identity holds:

$$\vec{a} \times (\vec{b} \times \vec{c}) = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c}. \qquad (2.90)$$

The mixed product $(\vec{a} \times \vec{b}) \cdot \vec{c}$ gives the volume of the parallelepiped spanned by $\vec{a}$, $\vec{b}$, and $\vec{c}$.

2.4.3 Angular Acceleration

The acceleration is calculated from the time derivative of the velocity $\vec{v}$:

$$\vec{a} = \frac{d\vec{v}}{dt} = \frac{d}{dt}(r\omega\vec{e}_T) = r\omega^2\vec{e}_N + \dot{\omega}r\vec{e}_T = \vec{a}_N + \vec{a}_T \tag{2.91}$$

with $d/dt\ \vec{e}_T = \omega\vec{e}_N$, where $\vec{e}_N$ points to the center of the circle. For

$$\vec{v} = \vec{\omega} \times \vec{r} \tag{2.92}$$

follows:

$$\vec{a} = \frac{d\vec{\omega}}{dt} \times \vec{r} + \vec{\omega} \times \vec{v}. \tag{2.93}$$

The component

$$\vec{a}_T = \frac{d\vec{\omega}}{dt} \times \vec{r} \tag{2.94}$$

is the **tangential component** of $\vec{a}$, to which the **normal component** of $\vec{a}$,

$$\vec{a}_N = \vec{\omega} \times \vec{v} = \vec{\omega} \times (\vec{\omega} \times \vec{r}), \tag{2.95}$$

is orthogonal.

Special case: uniform circular motion ($\vec{\omega} = \text{const}$) : With $\dot{\omega} = 0$ and $\vec{r} = -r\vec{e}_N$ follows

$$\vec{a} = \vec{a}_N = \vec{\omega} \times (\vec{\omega} \times \vec{r}) = (\vec{\omega} \cdot \vec{r})\vec{\omega} - (\vec{\omega} \cdot \vec{\omega})\vec{r} = r\omega^2\vec{e}_N \ : \text{centripetal acceleration} \tag{2.96}$$

Example: Motion of a mass point fixed on the earth's surface Fig. 2.5.

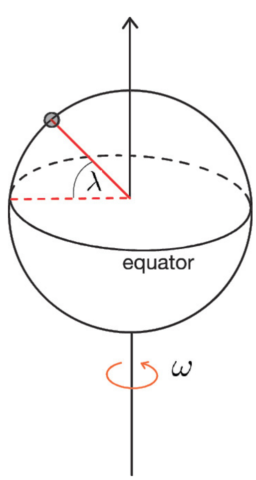

Fig. 2.5 Position of a mass point fixed on the earth's surface

The following applies to the velocity of the mass point:

$$v = \omega R \sin(90^\circ - \lambda) = \omega R \cos(\lambda) \quad , \tag{2.97}$$

where R is the Earth's radius, ω the magnitude of the angular velocity and λ the geographical width. The acceleration $\overrightarrow{a} = \overrightarrow{a}_N$ with magnitude

$$|\overrightarrow{a}_N| = \omega v = \omega^2 R \cos \lambda \tag{2.98}$$

points to the center of the circular path of the considered mass point; it is perpendicular to the north-south axis of the earth and to the velocity $\overrightarrow{v}$, which is directed tangential to the circular path.

Finally, we define the **angular acceleration** by:

$$\boxed{\overrightarrow{\alpha} = \frac{d\overrightarrow{\omega}}{dt}, \tag{2.99}}$$

which is the change in angular velocity over time.

In summarizing this chapter, we have described the motion of mass points in space and time and their trajectories by vectors $\overrightarrow{r}(t)$, $\overrightarrow{v}(t)$ and $\overrightarrow{a}(t)$. The mathematical tools are differentiation in cartesian or polar coordinate systems of a three-dimensional real vector space, which is used to uniquely define physical quantities like inertial systems, velocity, acceleration, angular velocity or angular acceleration. A brief introduction to Euclidean vector spaces has been given and linear transformations (like rotations) been described by suitable 3×3 matrices. This allows for a rigid formulation of kinematics also in case of circular motion.

Footnotes

1 $\mathbb{R}^3$ is the three dimensional real vector space.

W. Cassing, *Theoretical Physics Compact I*
https://doi.org/10.1007/978-3-031-95512-9_3

3. Relative Motion

Wolfgang Cassing[1]
(1) University of Giessen, Gießen, Hessen, Germany

Once the physical quantities of interest are defined it remains to clarify the conditions, that observers in different inertial systems—moving with some constant velocity $\vec{v}_0$ relative to each other—can find their observations to be identical. This leads us to the **Galilean relativity principle** and the **Galilei group of transformations**. Of special interest are rotating and center of mass systems, which will be discussed in detail.

3.1 Inertial Systems

3.1.1 Idea and Practice

According to Newton's first axiom (Sect. 1.2) an **inertial system** is defined by a uniform motion of a **free** particle. The practical use of Newton's axioms therefore depends on the question, if there are (at least approximately) inertial systems.

In the following we want to start from the idealized assumption that a strict inertial system was found. In this system the 2nd Newton axiom applies in the form

$$m\vec{a} = \vec{F}, \tag{3.1}$$

where the mass m is viewed as a positive constant. Adding the principle of Actio = Reactio (3rd axiom),

$$\vec{F}_{12} = -\vec{F}_{21}, \tag{3.2}$$

we can consider

$$\boxed{m\vec{a} = \vec{F} \qquad (3.3)}$$

as the definition of **force** and from the combination of the two equations obtain a rule for the measurement of mass.

3.1.2 Galilean Principle of Relativity

In addition to an inertial system Σ we consider another reference system Σ', which moves with constant velocity $\overrightarrow{v_0}$ relative to Σ. A mass point $\overrightarrow{P}$, whose position in the system Σ is given by the position vector $\overrightarrow{r}$, is described by the position vector $\overrightarrow{r}'$ in Σ':

$$\boxed{\overrightarrow{r}' = \overrightarrow{r} - \overrightarrow{v_0}t.} \tag{3.4}$$

Taking time derivatives we get for the velocities

$$\boxed{\overrightarrow{v}' = \overrightarrow{v} - \overrightarrow{v_0}} \tag{3.5}$$

and for the accelerations

$$\boxed{\overrightarrow{a}' = \overrightarrow{a}.} \tag{3.6}$$

If *P* moves freely in the system Σ, then *P* also moves freely with respect to Σ'. An observer in Σ' comes to the same result for the force as in Σ,

$$\boxed{\overrightarrow{F}' = m'\overrightarrow{a}' = m\overrightarrow{a} = \overrightarrow{F}.} \tag{3.7}$$

This identity finds its expression in the **Galilean principle of relativity**

> The basic laws of mechanics are the same in all reference systems moving relative to each other with a constant velocity.

With the assumption that time measurements are the same in all inertial frames,

$$t = t' \tag{3.8}$$

we define a **Galilean transformation** by:

$$\boxed{\overrightarrow{r}' = \overrightarrow{r} - \overrightarrow{v_0}t; \quad t' = t.} \tag{3.9}$$

Galilei's addition law for velocities then gives:

$$\vec{v}' = \vec{v} - \vec{v}_0. \tag{3.10}$$

Limits of Galilei's principle of relativity:

1. The basic equations of electrodynamics are not invariant with respect to the transformation $\vec{v}' = \vec{v} - \vec{v}_0$.
2. For high velocities ($v \leq c$; c: velocity of light) Newton's equation of motion is no longer applicable.

3.1.3 Galilei Group

The Galilean transformations form a **commutative group** $G(\vec{v}_0)$, where the connection between group elements is the successive execution of transformations.

1. **Commutativity**: The combination of two Galilei transformations gives a Galilei-transformation and is commutative.
2. **Associativity**: The associativity of the Galilean transformations follows from the associativity of the addition of velocities.
3. **Neutral element**: There is a neutral element,

$$\vec{v}_0 = \vec{0} \tag{3.11}$$

describing the identical transformation.

4. **Inverse element**: For every Galilean transformation G, characterized by the relative velocity $\vec{v}_0$ of the systems under consideration, there is an inverse Galilei-transformation G^{-1}, which corresponds to the relative velocity $-\vec{v}_0$, i.e. $G^{-1}(\vec{v}_0) = G(-\vec{v}_0)$.

3.2 Rotating Reference Systems

3.2.1 Non-inertial Systems

In **inertial systems** the equation of motion applies in the simple form:

$$m\vec{a} = \vec{F}. \tag{3.12}$$

However, from time to time it may be useful to switch to a **non-inertial system** in which the trajectory has a **simpler** form. To do this we have to know how velocity and acceleration change in the transformation from the inertial system to the non-inertial system.

3.2.2 Uniformly Rotating Systems

We consider the motion of a mass point in an inertial system Σ and in a system rotating uniformly relative to Σ, i.e. Σ'. Both systems should initially have the same origin. The position vector $\overrightarrow{r} \equiv \overrightarrow{r}'$ of the mass point is (at the initial time):

$$\overrightarrow{r} = x\overrightarrow{e_x} + y\overrightarrow{e_y} + z\overrightarrow{e_z} = x'\overrightarrow{e_{x'}} + y'\overrightarrow{e_{y'}} + z'\overrightarrow{e_{z'}} = \overrightarrow{r}'. \tag{3.13}$$

Here $\overrightarrow{e_i}$ and $\overrightarrow{e_{i'}}$ are unit orthogonal vectors in the direction of the cartesian axes of Σ or Σ'.

The velocity $\overrightarrow{v}$ for the observer in Σ at fixed coordinate system $\overrightarrow{e_i}$ is:

$$\boxed{\overrightarrow{v} = \frac{d\overrightarrow{r}}{dt} = v_x\overrightarrow{e_x} + v_y\overrightarrow{e_y} + v_z\overrightarrow{e_z}} \tag{3.14}$$

and for the observer in Σ' with fixed coordinate system $\overrightarrow{e_i}'$:

$$\boxed{\overrightarrow{v}' = \frac{d\overrightarrow{r}'}{dt} = v'_{x'}\overrightarrow{e_{x'}} + v'_{y'}\overrightarrow{e_{y'}} + v'_{z'}\overrightarrow{e_{z'}}.} \tag{3.15}$$

For the observer in Σ the axes of Σ' rotate; the vectors $\overrightarrow{e_{i'}}$ change in time, such that (in the system Σ) $\overrightarrow{v}$ can also be calculated as

$$\overrightarrow{v} = v'_{x'}\overrightarrow{e_{x'}} + x'\frac{d\overrightarrow{e_{x'}}}{dt} + v'_{y'}\overrightarrow{e_{y'}} + y'\frac{d\overrightarrow{e_{y'}}}{dt} + v'_{z'}\overrightarrow{e_{z'}} + z'\frac{d\overrightarrow{e_{z'}}}{dt}. \tag{3.16}$$

Then

$$\boxed{\overrightarrow{v} = \overrightarrow{v}' + \overrightarrow{\omega} \times \overrightarrow{r}'} \tag{3.17}$$

and $\overrightarrow{\omega}$ is the angular velocity with which Σ' rotates relative to Σ.

The **acceleration** for an observer in Σ is given by:

$$\boxed{\overrightarrow{a} = \frac{d\overrightarrow{v}}{dt} = a_x\overrightarrow{e_x} + a_y\overrightarrow{e_y} + a_z\overrightarrow{e_z}} \tag{3.18}$$

and for an observer in Σ' by:

$$\vec{a}\,' = \frac{d\vec{v}\,'}{dt} = a'_{x'}\vec{e}_{x'} + a'_{y'}\vec{e}_{y'} + a'_{z'}\vec{e}_{z'}. \tag{3.19}$$

For the observer in Σ the vectors $\vec{e}_{i'}$ are time dependent; accordingly

$$\vec{a} = \frac{d}{dt}(v'_{x'}\vec{e}_{x'} + x'\frac{d\vec{e}_{x'}}{dt} + v'_{y'}\vec{e}_{y'} + y'\frac{d\vec{e}_{y'}}{dt} + v'_{z'}\vec{e}_{z'} + z'\frac{d\vec{e}_{z'}}{dt}) =$$

$$\vec{a}\,' + v'_{x'}\frac{d\vec{e}_{x'}}{dt} + v'_{y'}\frac{d\vec{e}_{y'}}{dt} + v'_{z'}\frac{d\vec{e}_{z'}}{dt} + (\vec{\omega}\times\vec{v}) + \vec{\omega}\times(x'\frac{d\vec{e}_{x'}}{dt} + y'\frac{d\vec{e}_{y'}}{dt} + z'\frac{d\vec{e}_{z'}}{dt}) =$$

$$\vec{a}\,' + 2(\vec{\omega}\times\vec{v}\,') + \vec{\omega}\times(\vec{\omega}\times\vec{r}\,'). \tag{3.20}$$

The term $2(\vec{\omega}\times\vec{v}\,')$ is the **Coriolis acceleration** and the term $\vec{\omega}\times(\vec{\omega}\times\vec{r}\,')$ the **centrifugal acceleration**.

The equation of motion in the rotating system Σ' results from $\vec{F} = m\vec{a}$ in Σ:

$$m\vec{a}\,' = \vec{F} - 2\,m(\vec{\omega}\times\vec{v}\,') - m\vec{\omega}\times(\vec{\omega}\times\vec{r}\,'). \tag{3.21}$$

Thus in Σ', in addition to Newton's force $\vec{F}$, so-called **inertial forces** show up:

the **Coriolis force**

$$-2\,m(\vec{\omega}\times\vec{v}\,') \tag{3.22}$$

and the **centrifugal force**

$$-m\vec{\omega}\times(\vec{\omega}\times\vec{r}\,'). \tag{3.23}$$

In contrast to the forces determined by Newton's 2nd axiom the inertial forces do not contribute to the Newtonian forces for the interaction between mass points.

3.2.3 Explanations and Examples

A mass point is located at the end of a stretched thread on a circular path moving with constant angular velocity $\vec{\omega}$.

1. From the perspective of an observer in the inertial system Σ a force acts on the particle via the stretched thread,

$$\vec{F} = m\vec{a} = m\vec{\omega} \times (\vec{\omega} \times \vec{r}), \tag{3.24}$$

which accelerates the particle towards the center of the circle.

2. For an observer in the (co-)rotating system Σ' the particle is not accelerated; $\vec{a}\,' = 0$. This can be interpreted as follows: in Σ' the centrifugal force and the Newtonian force $\vec{F}$, originating from the stretched thread, balance each other.

3.2.4 Generalization

In case that the origin of Σ' is not the same as that of Σ, i.e. $\vec{r} = \vec{R} + \vec{r}\,'$, we get

$$\boxed{\vec{v} = \dot{\vec{R}} + \vec{v}\,' + (\vec{\omega} \times \vec{r}\,')} \tag{3.25}$$

and

$$\boxed{\vec{a} = \ddot{\vec{R}} + \vec{a}' + 2(\vec{\omega} \times \vec{v}\,') + \vec{\omega} \times (\vec{\omega} \times \vec{r}\,,'\,),} \tag{3.26}$$

if Σ' is accelerated relative to Σ (by $\ddot{\vec{R}}$) or is moving with relative velocity $\dot{\vec{R}}(t)$.

3.3 Center of Mass System

3.3.1 Definition of the Center of Mass

$$\boxed{\vec{r}_s = \tfrac{1}{M} \sum_{i=1}^{N} m_i \vec{r}_i \quad ; \quad M = \sum_{i=1}^{N} m_i \quad (\text{total mass})\ ,} \tag{3.27}$$

where m_i are the particle masses and $\vec{r}_i$ their positions in a space-fixed coordinate system Σ. We obtain for the velocity of the center of mass of N particles:

$$\overrightarrow{v}_s = \frac{1}{M} \sum_{i=1}^{N} m_i \overrightarrow{v}_i \tag{3.28}$$

and for the acceleration:

$$\overrightarrow{a}_s = \frac{1}{M} \sum_{i=1}^{N} m_i \overrightarrow{a}_i. \tag{3.29}$$

If the system is an inertial system, then according to Newton's 2nd axiom:

$$m_i \overrightarrow{a}_i = \overrightarrow{F}_i, \quad i = 1, 2, \ldots, N. \tag{3.30}$$

The equation of motion for the center of mass then is given by Newton's 4th axiom:

$$M\overrightarrow{a}_s = \overrightarrow{F}_s \text{ with } \overrightarrow{F}_s = \sum_{i=1}^{N} \overrightarrow{F}_i. \tag{3.31}$$

If no **external forces** $\overrightarrow{F}_s$ act, we get

$$\overrightarrow{F}_s = 0 \tag{3.32}$$

(according to Newton's 3rd axiom), since the internal forces between the particles cancel in pairs, i.e.

$$M\overrightarrow{a}_s = 0\,; \tag{3.33}$$

the center of mass moves uniformly on a straight line.

3.3.2 Observables in the Center of Mass System

For many problems it is useful to move from the laboratory system to the center of mass system. We consider a **closed** system which is defined by vanishing external forces. We now move to the center of mass system Σ' by the condition that the center of mass is at rest:

$$\overrightarrow{v}_s{}' = 0. \tag{3.34}$$

If one specifically chooses the center of mass as the origin of the system Σ', we have

$$\overrightarrow{r_s}\,' = 0. \tag{3.35}$$

The positions of the particles then are:

$$\overrightarrow{r_i}\,' = \overrightarrow{r_i} - \overrightarrow{r_s}; \text{accordingly} \sum_i m_i \overrightarrow{r_i}\,' = 0. \tag{3.36}$$

The velocities in Σ' are:

$$\overrightarrow{v_i}\,' = \overrightarrow{v_i} - \overrightarrow{v_s}\,, \tag{3.37}$$

and the accelerations:

$$\overrightarrow{a_i}\,' = \overrightarrow{a_i} - \overrightarrow{a_s}\,. \tag{3.38}$$

Thus:

$$\boxed{\sum_{i=1}^{N} m_i \overrightarrow{v_i}\,' = 0.} \tag{3.39}$$

3.3.3 Determination of the Center of Mass

In case of a continuous mass distribution $\rho(x, y, z)$ the center of mass vector is given by:

$$\boxed{\overrightarrow{r_s} = \frac{1}{M} \iiint\limits_V \overrightarrow{r}\,\rho(x, y, z)\;dx\,dy\,dz,} \tag{3.40}$$

with the total mass:

$$\boxed{M = \iiint\limits_V \rho(x, y, z)\;dx\,dy\,dz,} \tag{3.41}$$

i.e. the summation over mass points m_i at positions $\overrightarrow{r_i}$ is replaced by an integration over space with the mass distribution $\rho(\overrightarrow{r})$.

3.3.4 Collision of Two Particles

In the system Σ (inertial system) we have:

$$m_1\overrightarrow{a_1} = \overrightarrow{F}_{12} \quad ; \quad m_2\overrightarrow{a_2} = \overrightarrow{F}_{21} = -\overrightarrow{F}_{12} \quad , \tag{3.42}$$

if no external forces are at work. In the center of mass system we then get (see Fig. 3.1):

$$m_1\overrightarrow{v_1}\,' = -m_2\overrightarrow{v_2}\,' \tag{3.43}$$

both before (left) and after the collision (right).

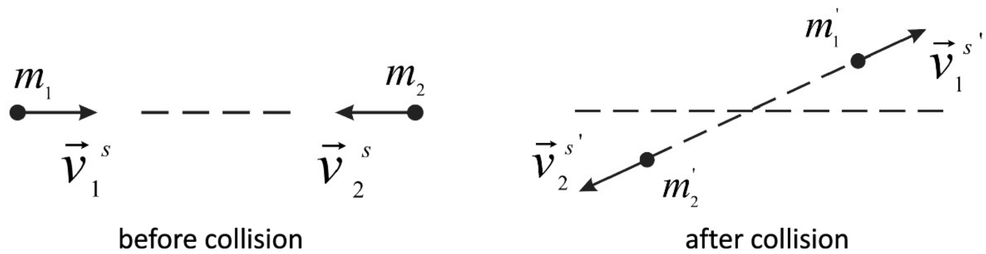

Fig. 3.1 Velocity vectors and positions before (left) and after the collision (right)

3.3.5 Reduced Mass

The advantage of the center of mass system is that the number of degrees of freedom is reduced. After separating the center of mass motion only 3 degrees of freedom remain for the 2-particle problem! By introducing the relative vector

$$\overrightarrow{r} = \overrightarrow{r_1}^{\,s} - \overrightarrow{r_2}^{\,s} = \overrightarrow{r_1} - \overrightarrow{r_2} \tag{3.44}$$

the equation of motion for the relative motion reads:

$$\boxed{\mu\ddot{\overrightarrow{r}} = \mu(\ddot{\overrightarrow{r_1}} - \ddot{\overrightarrow{r_2}}) = \mu(\tfrac{\overrightarrow{F_{12}}}{m_1} - \tfrac{\overrightarrow{F_{21}}}{m_2}) = \mu(\tfrac{1}{m_1} + \tfrac{1}{m_2})\overrightarrow{F}_{12} = \mu\tfrac{m_1+m_2}{m_1m_2}\overrightarrow{F}_{12} = \overrightarrow{F}_{12} \quad (3.45)}$$

with the **reduced mass**

$$\boxed{\mu = \tfrac{m_1m_2}{m_1+m_2}. \qquad (3.46)}$$

Since the problem of the center of mass motion has already been solved (in the absence of external forces), we have reduced the two-body problem (6 degrees of freedom) to an effective one-body problem (3 degrees of freedom).

Two **simple limiting cases** are:

1. $m_1 = m_2 = m$. Then we get

$$\mu = \tfrac{1}{2}m \tag{3.47}$$

for example in proton-proton scattering.

2. $m_1 \gg m_2$. In this case we obtain:

$$\mu = \frac{m_2}{1+\frac{m_2}{m_1}} \approx m_2. \qquad (3.48)$$

Thus the mass of the lighter particle is approximately giving the reduced mass, e.g. for the motion of an electron around a nucleus or of the earth around the sun.

In summarizing this chapter we have specified inertial systems and introduced the Galilean principle of relativity to compare observations in different inertial systems, which move relative to each other with a constant velocity $\overrightarrow{v_0}$. Of special interest have been rotating systems and center of mass systems, where the latter is of particular importance for the description of binary collisions.

W. Cassing, *Theoretical Physics Compact I*
https://doi.org/10.1007/978-3-031-95512-9_4

4. Dynamics

Wolfgang Cassing[1]
(1) University of Giessen, Gießen, Hessen, Germany

After the preparatory work in the previous chapters we here will define forces and derive Newton's equations of motion; their solution will provide the trajectory of a mass point in space and time. Examples for characteristic problems will be given and the explicit solutions derived in detail. It will turn out that instead of velocities or angular velocities it is more convenient to introduce momenta and angular momenta of particles since for closed systems—without external forces—the total momentum is a constant of motion. This also holds for the angular momentum if no external torque acts on the system. Next we will consider the connection between the work done by a force on a particle along its trajectory and the actual kinetic energy. In case of conservative forces we can introduce a potential energy $U(\vec{r})$ that allows to compute the actual force by its negative gradient. Then the energy of the system can be defined by the sum of kinetic and potential energy and —for closed systems—is found to be a conserved quantity, too.

4.1 Consequences from Newton's Axioms

The explicit formulation of Newton's axioms has been given in Sect. 1.2.

4.1.1 Mass

The combination of the 2nd and 3rd axioms (in Sect. 1.2) for the collision of 2 particles with masses m_1 and m_2 leads to:

$$\tfrac{d}{dt}(m_1\vec{v}_1) = \vec{F}_{12} \tag{4.1}$$

$$\tfrac{d}{dt}(m_2\vec{v}_2) = \vec{F}_{21} = -\vec{F}_{12}. \tag{4.2}$$

It follows:

(4.3)

$$\frac{d}{dt}(\overrightarrow{p_1} + \overrightarrow{p_2}) = 0$$

for the **momenta** of the particles $(i = 1, 2)$ defined by

$$\boxed{\overrightarrow{p_i} = m_i \overrightarrow{v_i}.} \tag{4.4}$$

The sum of the momenta in the collision is thus constant in time:

$$\overrightarrow{p_1}' - \overrightarrow{p_1} = \Delta\overrightarrow{p_1} = -\Delta\overrightarrow{p_2} = -(\overrightarrow{p_2}' - \overrightarrow{p_2}) \tag{4.5}$$

or

$$\boxed{\frac{m_2}{m_1} = \frac{|\Delta\overrightarrow{v_1}|}{|\Delta\overrightarrow{v_2}|} \quad ,} \tag{4.6}$$

if the mass is independent of the body’s state of motion.

We can use Eq. (4.6) as an operational definition of **mass**: We can do this by measuring velocities and thus determine the ratio of two masses, i.e. if one mass is specified, but chosen uniquely as **unit mass** m_1, the mass m_2 relative to m_1 is fixed. The question of whether the mass possibly might be velocity-dependent can be determined by scattering experiments: One finds that in non-relativistic mechanics $(v \ll c)$ the mass can be assumed to be independent of the velocity.

4.1.2 Force

Since we have introduced the mass as a scalar, the force is (in line with the 2nd axiom)—like the acceleration—a vector:

$$\overrightarrow{F} = m\overrightarrow{a}. \tag{4.7}$$

The superposition principle (4th axiom) does not follow from the vector character of the force

$$\overrightarrow{F} = \overrightarrow{F_1} + \overrightarrow{F_2}, \tag{4.8}$$

because the vector property of the force would also be satisfied if for the resulting force we have

(4.9)

$$\overrightarrow{F} = \overrightarrow{F_1} + \overrightarrow{F_2} + \overrightarrow{f}(\overrightarrow{F_1}, \overrightarrow{F_2}).$$

The vector function $\overrightarrow{f}$ is introduced here to account for a possible mutual influence of the forces $\overrightarrow{F_1}$ and $\overrightarrow{F_2}$. The superposition principle is therefore an **independent** axiom, which not automatically follows from the vector character of the force!

4.1.3 Equations of Motion

The equations of motion for a system of *N* mass points are:

$$\boxed{m_i \overrightarrow{a_i} = \overrightarrow{F_i}, \qquad i = 1, 2, 3, \ldots, N} \tag{4.10}$$

where $\overrightarrow{F_i}$ is the total force acting on particle i. It is composed additively by

1. **internal forces**,
   from the interaction with the $(N-1)$ particles, for which the 3rd axiom applies,
2. **external forces**,
   describing the influence of the environment.

Mathematically speaking, the equations of motion are generally a coupled system of 2nd order differential equations for the trajectories $\overrightarrow{r_i}(t)$ that have to be calculated. One obtains unique solutions if the initial conditions

$$\overrightarrow{r_i}(t_0) = \overrightarrow{r_i}^{\,0} \tag{4.11}$$

$$\overrightarrow{v_i}(t_0) = \overrightarrow{v_i}^{\,0} \tag{4.12}$$

are known at some time t_0. These are $2 \cdot 3 \cdot N = 6N$ boundary conditions.

Example: Motion of a particle in a single dimension:

From the equation of motion

$$m\ddot{x} = F(t) \tag{4.13}$$

we obtain by integration in time

$$\dot{x}(t) = \frac{1}{m} \int_{t_0}^{t} F(t') \; dt' + c_1 \tag{4.14}$$

and by further integration

$$x(t) = \int_{t_0}^{t} \dot{x}(t')dt' + c_2. \tag{4.15}$$

The two integration constants c_1 and c_2 are determined as soon as the initial conditions at t_0 are known:

$$\dot{x}(t_0) = c_1 \qquad x(t_0) = c_2. \tag{4.16}$$

4.2 Examples for Solving Equations of Motion

4.2.1 Charged Particle in a Homogeneous Electric Field

The force on a point charge q in an electrostatic field is given by

$$\overrightarrow{F} = q\overrightarrow{E}, \tag{4.17}$$

where $\overrightarrow{E}$ is the electric field strength, which we take as spatial and assume to be constant in time, i.e. $\overrightarrow{E} = \overrightarrow{E}(\overrightarrow{r})$.

The equation of motion then reads:

$$m\overrightarrow{a} = q\overrightarrow{E}. \tag{4.18}$$

Let's choose the coordinate system such that

$$\overrightarrow{E} = \begin{pmatrix} 0 \\ 0 \\ E_z \end{pmatrix}. \tag{4.19}$$

In this case the equations of motion simplify to:

$$\ddot{x} = 0 \qquad \ddot{y} = 0 \qquad \ddot{z} = \frac{q}{m} E_z \quad . \tag{4.20}$$

By time integration we get ($t_0 = 0$)

$$\dot{x} = v_x(0) \qquad \dot{y} = v_y(0) \qquad \dot{z} = \frac{qE_z}{m} t + v_z(0) \tag{4.21}$$

for the velocities. Repeated integration gives

$$x = x_0 + v_x(0)t \qquad y = y_0 + v_y(0)t \qquad z = z_0 + v_z(0)t + \frac{qE_z}{2\,m} t^2. \tag{4.22}$$

In vector notation:

$$\overrightarrow{r}(t) = \overrightarrow{r_0} + \overrightarrow{v}(0)t + \tfrac{q}{2\,m}\overrightarrow{E}t^2. \tag{4.23}$$

Important special cases:

1. $\overrightarrow{v}(0)$ parallel $\overrightarrow{E}$. We obtain

$$x, y = \text{const.} \qquad \mathrm{z = z_0 + v_z(0)t + \tfrac{q}{2\,m}E_z t^2}. \tag{4.24}$$

   There is a rectilinear accelerated motion as in case of the free fall.

2. $\overrightarrow{v}(0)$ perpendicular to $\overrightarrow{E}$.

   With a suitable choice of coordinates we get for $\overrightarrow{v}(0) = v_y(0)\overrightarrow{e_y}$

$$x(t) = 0 \qquad y(t) = v_y(0)t \qquad z(t) = \tfrac{qE_z}{2\,m}t^2; \tag{4.25}$$

   a parabola results for the trajectory:

$$z(t) = \tfrac{qE_z}{2mv_y^2(0)}y^2(t). \tag{4.26}$$

4.2.2 Charged Particle in a Constant Homogeneous Magnetic Field

The force on a particle with charge *q* and velocity $\overrightarrow{v}$ in a magnetic field $\overrightarrow{B}$ is given by:

$$\overrightarrow{F} = \tfrac{q}{c}(\overrightarrow{v} \times \overrightarrow{B}) \quad (c: \text{velocity of light}) \quad . \tag{4.27}$$

Let’s choose the coordinate system such that

$$\overrightarrow{B} = \begin{pmatrix} 0 \\ 0 \\ B_z \end{pmatrix}, \tag{4.28}$$

which leads to:

$$\overrightarrow{v} \times \overrightarrow{B} = \begin{pmatrix} v_y B_z \\ -v_x B_z \\ 0 \end{pmatrix}. \tag{4.29}$$

The equation of motion then reads:

$$a_x = \frac{q}{mc} v_y B_z \qquad a_y = -\frac{q}{mc} v_x B_z \qquad a_z = 0. \tag{4.30}$$

Obviously $\overrightarrow{a}$ is perpendicular to $\overrightarrow{v}$,

$$\overrightarrow{v} \cdot \overrightarrow{a} = 0, \tag{4.31}$$

such that

$$\frac{d}{dt} v^2 = \frac{d}{dt} (\overrightarrow{v} \cdot \overrightarrow{v}) = 2\overrightarrow{v} \cdot \overrightarrow{a} = 0 \tag{4.32}$$

or

$$v^2 = \text{const}. \tag{4.33}$$

In *z*-direction the motion is trivial:

$$v_z = \text{const, thus : z(t) = z}_0 + \text{v}_\text{z}(0)\text{t}. \tag{4.34}$$

The equations of motion are coupled in the *x*, *y* direction. To find the solution, we first use the complex auxiliary variable

$$Q(t) = x(t) + iy(t). \tag{4.35}$$

Differentiation with respect to *t* yields

$$\dot{Q} = \dot{x} + i\dot{y} = v_x + iv_y \tag{4.36}$$

and

$$\ddot{Q} = \ddot{x} + i\ddot{y} = a_x + ia_y. \tag{4.37}$$

For the variable $a_x + ia_y$ we then obtain

$$a_x + ia_y = \frac{qB_z}{mc} (v_y - iv_x) \tag{4.38}$$

or

$$\ddot{Q} = -i \frac{qB_z}{mc} \dot{Q}. \tag{4.39}$$

With the **Ansatz**,

$$Q = Q_0 e^{\lambda t}, \tag{4.40}$$

we get by insertion:

$$\lambda^2 = -i\omega\lambda \quad \text{with } \omega = \tfrac{qB_z}{mc}, \tag{4.41}$$

thus

$$\lambda = 0 \text{ or } \lambda = -i\omega. \tag{4.42}$$

The general solution then reads:

$$Q = Q_{01} + Q_{02} e^{-i\omega t}. \tag{4.43}$$

The 2 complex constants Q_{01} and Q_{02} are determined by the 4 real initial conditions for $x(0), y(0), \dot{x}(0)$ and $\dot{y}(0)$:

$$x(0) + iy(0) = Q(0) = Q_{01} + Q_{02} \tag{4.44}$$

$$\dot{x}(0) + i\dot{y}(0) = -i\omega Q_{02}. \tag{4.45}$$

Writing Q_{02} as

$$Q_{02} = \varrho e^{i\alpha} = (\rho \cos \alpha + i\rho \sin \alpha), \tag{4.46}$$

i.e. in polar coordinates, we get:

$$\dot{x}(0)^2 + \dot{y}(0)^2 = v_\perp^2 = \omega^2 \varrho^2. \tag{4.47}$$

Thus

$$\varrho = v_\perp / \omega, \tag{4.48}$$

where $v_\perp$ is the magnitude of the velocity perpendicular to the z-direction. For the phase α one finds:

$$\tan \alpha = -\tfrac{\dot{x}(0)}{\dot{y}(0)}. \tag{4.49}$$

Dividing $Q(t)$ (4.43) again into real and imaginary parts, we obtain:

$$x = x_0 + \varrho \cos (\alpha - \omega t), \tag{4.50}$$

$$\tag{4.51}$$

$$y = y_0 + \varrho \sin\,(\alpha - \omega t)$$

with $x_0 = x(0) - \rho \cos\,\alpha$ and $y_0 = y(0) - \rho \sin\,\alpha$. The trajectory then describes a circle

$$(x - x_0)^2 + (y - y_0)^2 = \varrho^2 \tag{4.52}$$

with radius ϱ and center (x_0, y_0).

4.2.3 Free Fall on the Rotating Earth

Approximate inertial system:

We choose a system Σ' whose origin is at the center of the earth and whose axis directions are firmly defined relative to the fixed stars. In Σ' then (approximately) holds:

$$m\overrightarrow{a}' = \overrightarrow{F}, \tag{4.53}$$

where $\overrightarrow{F}$ is the gravitational force between the mass point of mass *m* and the earth.

We now move to a system Σ that rotates rigidly with the earth, whose origin is located on the earth's surface. Then—according to Sect. 3.2.4—(exchanging Σ and Σ') we get:

$$\overrightarrow{a}' = \ddot{\overrightarrow{R}} + \overrightarrow{a} + 2(\overrightarrow{\omega} \times \overrightarrow{v}) + \overrightarrow{\omega} \times (\overrightarrow{\omega} \times \overrightarrow{r}), \tag{4.54}$$

where $\overrightarrow{R}$ is the vector from the center of the earth Σ' to the origin of the system Σ rotating with the earth. Σ moves in a circular path with (constant) angular velocity $\overrightarrow{\omega}$ of the earth's rotation. Therefore:

$$\ddot{\overrightarrow{R}} = \overrightarrow{\omega} \times (\overrightarrow{\omega} \times \overrightarrow{R}), \tag{4.55}$$

such that

$$\overrightarrow{a} = \overrightarrow{g}(\lambda) - 2(\overrightarrow{\omega} \times \overrightarrow{v}) - \overrightarrow{\omega} \times (\overrightarrow{\omega} \times \overrightarrow{r}), \tag{4.56}$$

where

$$\overrightarrow{g}(\lambda) = \frac{\overrightarrow{F}}{m} - \overrightarrow{\omega} \times (\overrightarrow{\omega} \times \overrightarrow{R}) \tag{4.57}$$

is the '**effective**' gravitational acceleration in Σ.

We now make the approximation that the height of the fall is small compared to the distance R of the systems Σ and Σ', i.e. $|\overrightarrow{r}| \ll |\overrightarrow{R}|$. Then we can neglect $\overrightarrow{\omega} \times (\overrightarrow{\omega} \times \overrightarrow{r})$ relative to $\overrightarrow{\omega} \times (\overrightarrow{\omega} \times \overrightarrow{R})$. The axes of the system Σ we define as follows: the z-axis is antiparallel to the effective gravitational acceleration $\overrightarrow{g}(\lambda)$, the x-axis in the north-south direction, and the y-axis in the west-east direction. Then we obtain:

$$\overrightarrow{\omega} = -\omega \sin \gamma \overrightarrow{e_x} + \omega \cos \gamma \overrightarrow{e_z}, \tag{4.58}$$

where γ is the angle between $\overrightarrow{e_z}$ and $\overrightarrow{e_{z'}}$. The equations of motion read:

$$\ddot{x} = 2\dot{y}\omega \cos \gamma \tag{4.59}$$

$$\ddot{y} = -2\dot{z}\omega \sin \gamma - 2\dot{x}\omega \cos \gamma \tag{4.60}$$

$$\ddot{z} = -g(\lambda) + 2\dot{y}\omega \sin \gamma. \tag{4.61}$$

As initial conditions we choose:

$$x(0) = 0 \qquad \dot{x}(0) = 0 \tag{4.62}$$

$$y(0) = 0 \qquad \dot{y}(0) = 0 \tag{4.63}$$

$$z(0) = z_0 \qquad \dot{z}(0) = 0. \tag{4.64}$$

Since the Coriolis force is a small correction to gravity we can write the solution as a Taylor series with respect to ω:

$$x = x_1 + \omega x_2 + \cdots \tag{4.65}$$

$$y = y_1 + \omega y_2 + \cdots \tag{4.66}$$

$$z = z_1 + \omega z_2 + \cdots \tag{4.67}$$

This approximation is inserted into the equations of motion (4.59), (4.60), (4.61) and one has to take care that they must be identical in ω. One gets:

$$\ddot{x}_1 = 0 \qquad \ddot{y}_1 = 0 \qquad \ddot{z}_1 = -g(\lambda), \tag{4.68}$$

and

$$\ddot{x}_2 = 2\dot{y}_1 \cos\gamma \tag{4.69}$$

$$\ddot{y}_2 = -2\dot{z}_1 \sin\gamma - 2\dot{x}_1 \cos\gamma \tag{4.70}$$

$$\ddot{z}_2 = 2\dot{y}_1 \sin\gamma \tag{4.71}$$

for the terms linear in ω. This leads to:

$$x_1 = 0 \qquad y_1 = 0 \qquad z_1 = z_0 - \tfrac{1}{2}g(\lambda)t^2 \tag{4.72}$$

and

$$\ddot{x}_2 = 0 \tag{4.73}$$

$$\ddot{y}_2 = 2gt \sin\gamma \tag{4.74}$$

$$\ddot{z}_2 = 0. \tag{4.75}$$

A special solution is:

$$x_2 = 0 \qquad y_2 = \tfrac{1}{3}gt^3 \sin\gamma \qquad z_2 = 0. \tag{4.76}$$

The complete solution then is:

$$x = 0 \qquad y = \tfrac{\omega}{3}gt^3 \sin\gamma \qquad z = z_0 - \tfrac{1}{2}gt^2. \tag{4.77}$$

Thus one gets an **east deviation** from the normal fall law.

Estimate: For $\gamma = 45^\circ$ and $z_0 = 100m$ the deviation is $y \approx 1.5$ cm; the effect is maximum at the equator.

4.3 Momentum and Angular Momentum

4.3.1 Momentum

The momentum of a particle of mass *m* is defined as

$$\tag{4.78}$$

$$\overrightarrow{p} = m\overrightarrow{v},$$

if $\overrightarrow{v}$ is its velocity. Since m is a scalar and $\overrightarrow{v}$ is a vector, $\overrightarrow{p}$ is also a vector. The Newtonian equation of motion then reads:

$$\frac{d\overrightarrow{p}}{dt} = \overrightarrow{F}. \tag{4.79}$$

In words: the **force is identical to the temporal change in momentum**. If there is no force acting on a particle, the momentum of the particle is constant in time:

$$\frac{d\overrightarrow{p}}{dt} = \overrightarrow{0} \rightarrow \overrightarrow{p} = \text{const.} \tag{4.80}$$

For a system of N particles with masses m_i the momentum of the i-th particle is given by:

$$\overrightarrow{p_i} = m_i\overrightarrow{v_i} \tag{4.81}$$

It's equation of motion is:

$$\frac{d\overrightarrow{p_i}}{dt} = \overrightarrow{F_i}\,, \tag{4.82}$$

where $\overrightarrow{F_i}$ is the total force acting on the i-th particle.

The total momentum of the N particles

$$\overrightarrow{P} = \sum_{i=1}^{N}\overrightarrow{p_i} = M\overrightarrow{v_s} \tag{4.83}$$

is a conserved quantity for a closed system (constant of motion).

The following holds:

$$\frac{d\overrightarrow{P}}{dt} = \sum_{i=1}^{N}\overrightarrow{F_i} = \overrightarrow{F_a}\,, \tag{4.84}$$

where $\overrightarrow{F_a}$ is the sum of all "external forces",

$$\overrightarrow{F_a} = \sum_{i=1}^{N} \overrightarrow{F_{ia}} \, . \tag{4.85}$$

The internal forces cancel each other in pairs because for each term $\overrightarrow{F_{ij}}$ also $\overrightarrow{F_{ji}} = -\overrightarrow{F_{ij}}$ in $\sum_i \overrightarrow{F_i}$ occurs. For a closed system the following holds:

$$\overrightarrow{F_{ia}} = 0, \text{thus}\, \overrightarrow{F_a} = 0 \, , \tag{4.86}$$

$$\frac{d\overrightarrow{P}}{dt} = 0 \;\rightarrow\; \overrightarrow{P} = \text{const.} \tag{4.87}$$

Thus the 3rd Newton axiom is crucial for the conservation of momentum in a closed system.

4.3.2 Momentum Law and Galilean Invariance

We assume that the momentum law holds in an inertial system Σ_v:

$$\sum_{i=1}^{N} m_i \overrightarrow{v_i} = \sum_{i=1}^{N'} m_i' \overrightarrow{v'}_i \, , \tag{4.88}$$

where m_i, $\overrightarrow{v_i}$ are the masses and velocities at some time t, m_i', $\overrightarrow{v'}_i$ at another time t'. By the distinction between m_i and m_i' as well as between N and N' we allow for mass exchange between the particles.

According to the principle of relativity the momentum law must also hold in every other inertialsystem Σ_u:

$$\sum_{i=1}^{N} m_i \overrightarrow{u_i} = \sum_{i=1}^{N'} m_i' \overrightarrow{u'}_i \, . \tag{4.89}$$

This gives the **law of mass conservation**:

$$M = \sum_{i=1}^{N} m_i = \sum_{i=1}^{N'} m_i' . \tag{4.90}$$

Proof If $\overrightarrow{v} \neq (0,0,0)$ is the velocity of the systems Σ_v and Σ_u relative to each other, then:

$$\overrightarrow{v_i} = \overrightarrow{u_i} + \overrightarrow{v}\,; \overrightarrow{v'}_i = \overrightarrow{u'}_i + \overrightarrow{v}. \tag{4.91}$$

This leads to (4.88):

$$\sum_{i=1}^{N} m_i \overrightarrow{u_i} + \overrightarrow{v} \sum_{i=1}^{N} m_i = \sum_{i=1}^{N'} m_i' \overrightarrow{u'}_i + \overrightarrow{v} \sum_{i=1}^{N'} m_i'\,, \tag{4.92}$$

and—due to momentum conservation – to:

$$\overrightarrow{0} = \overrightarrow{v}\Big(\sum_{i=1}^{N} m_i - \sum_{i=1}^{N'} m_{i'}\Big). \tag{4.93}$$

According to Galilei's principle of relativity the **momentum law** and **mass-conservation law** are connected to each other. (Note: This relationship does not hold in relativistic mechanics.)

4.3.3 Example: Rocket in Gravity-Free Space

We are looking for the velocity of the rocket as a function of the time changing mass according to the emission of mass Δm with velocity v_G. The momentum law holds because in gravity-free space there is no external force acting on the rocket.

We can then formulate the problem as follows: At time *t* the rocket has the mass $m = m(t)$ and the velocity $v = v(t)$ relative to the earth, which we want to consider as an inertial system. In the time Δt the rocket mass changes by $\Delta m < 0$; then the rocket at time $t + \Delta_t$ has the mass $(m + \Delta m)$ at a changed velocity $(v + \Delta\ v)$. The (positive) repelled amount of gas $(-\Delta\ m)$ has the velocity $(-v_G\ + v\ + \Delta\ v)$ relative to earth. According to the momentum law:

$$mv = (m + \Delta m)(v + \Delta v) + (-\Delta m)(-v_G + v + \Delta v) \tag{4.94}$$

or

$$0 = m\Delta v + \Delta m v_G\ . \tag{4.95}$$

The change of the velocity during the time Δt is:

$$\frac{\Delta v}{\Delta t} = -v_G \frac{1}{m} \frac{\Delta m}{\Delta t} \tag{4.96}$$

or in the limit $\Delta t \to 0$:

$$\frac{dv}{dt} = -v_G \frac{1}{m} \frac{dm}{dt}\,. \tag{4.97}$$

Integration in time gives:

$$v = v_0 + v_G \ln\left(\frac{m_0}{m}\right) , \tag{4.98}$$

if the rocket has mass m_0 and velocity v_0 at time t_0.

4.3.4 Angular Momentum

The angular momentum $\overrightarrow{l}$ of a particle with momentum $\overrightarrow{p}$ at position $\overrightarrow{r}$ is defined by

$$\boxed{\overrightarrow{l} = \overrightarrow{r} \times \overrightarrow{p} \quad .} \tag{4.99}$$

With $d\overrightarrow{p}/dt = \overrightarrow{F}$ we get

$$\boxed{\frac{d\overrightarrow{l}}{dt} = \overrightarrow{r} \times \overrightarrow{F} ,} \tag{4.100}$$

since $\dot{\overrightarrow{r}} \times \overrightarrow{p} = 0$, i.e. the temporal change of $\overrightarrow{l}$ is determined by the **torque**

$$\boxed{\overrightarrow{n} = \overrightarrow{r} \times \overrightarrow{F} .} \tag{4.101}$$

If there is no torque, $\overrightarrow{n} = 0$, the angular momentum is constant:

$$\boxed{\frac{d\overrightarrow{l}}{dt} = 0 \ \rightarrow \ \overrightarrow{l} = \text{const.}} \tag{4.102}$$

This is fulfilled for

1. $\overrightarrow{F} = 0$ trivially and for
2. **central forces**

(4.103)

$$\overrightarrow{F} = k(r)\overrightarrow{r},$$

such as the important cases of gravitational force or Coulomb force ($k(r) \sim 1/r^3$).

For N particles we define the total angular momentum as follows:

$$\overrightarrow{L} = \sum_{i=1}^{N} \overrightarrow{l_i} = \sum_{i=1}^{N} (\overrightarrow{r_i} \times \overrightarrow{p_i}). \tag{4.104}$$

The temporal change of the total angular momentum is:

$$\frac{d\overrightarrow{L}}{dt} = \sum_{i=1}^{N} (\overrightarrow{r_i} \times \overrightarrow{F_i}) = \sum_{i=1}^{N} \overrightarrow{n_i} = \overrightarrow{N}. \tag{4.105}$$

The angular momentum $\overrightarrow{L}$ therefore is constant in time if the total torque $\overrightarrow{N}$ vanishes.

4.3.5 Conservation of Angular Momentum and Galilean Invariance

We assume that the angular momentum of the system under consideration (in some inertial system Σ_v) is conserved. For the transition from Σ_v to another inertial system Σ_u we have:

$$\overrightarrow{r_i} \rightarrow \overrightarrow{r_i} - \overrightarrow{v}t\,; \overrightarrow{v_i} \rightarrow \overrightarrow{u_i} = \overrightarrow{v_i} - \overrightarrow{v}, \tag{4.106}$$

if $\overrightarrow{v}$ is the relative velocity between Σ_u and Σ_v. This leads to:

$$\sum_{i=1}^{N} (\overrightarrow{r_i} \times \overrightarrow{p_i}) \rightarrow \sum_{i=1}^{N} (\overrightarrow{r_i} - \overrightarrow{v}t) \times (\overrightarrow{p_i} - m_i\overrightarrow{v}) \tag{4.107}$$

or

$$\overrightarrow{L} \rightarrow \overrightarrow{L} + M\overrightarrow{v} \times (\overrightarrow{r_s} - \overrightarrow{v_s}t)\,, \tag{4.108}$$

where M is the total mass, $\overrightarrow{r_s}$ and $\overrightarrow{v_s}$ position and velocity of the center of mass. If no external force acts, the motion of the center of mass is a straight line, such that

(4.109)

$$\vec{r}_s - \vec{v}_s t = \text{const.}$$

$\Rightarrow$ The angular momentum $\vec{L}$ changes during the transition $\Sigma_v \to \Sigma_u$ only by an additive constant.

4.3.6 Examples

Uniform circular motion

The angular momentum $\vec{l} = \vec{r} \times \vec{p}$ is perpendicular to the plane of circular motion and has the magnitude

$$l = m\omega r^2 \,. \tag{4.110}$$

It is constant because for uniform circular motion ω and r are constant.

Area law

For a mass point under the influence of a central force the conservation of angular momentum implies

1. that the motion takes place in a plane F spanned by $\vec{r}$ and $\vec{v}$, and

2. that the **surface velocity** is constant,

$$\frac{d\vec{F}}{dt} = \text{const.} \tag{4.111}$$

Proof We consider the plane F spanned by the 2 neighboring position vectors $\vec{r}$ and $\vec{r} + \Delta\vec{r}$ (Fig. 4.1):

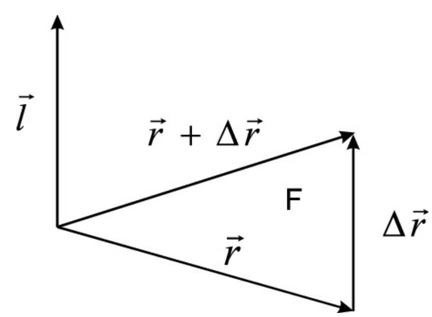

Fig. 4.1 Plane F spanned by the 2 neighboring position vectors $\vec{r}$ and $\vec{r} + \Delta\vec{r}$

$$\Delta\vec{F} = \frac{1}{2}\left(\vec{r} \times \left[\vec{r} + \Delta\vec{r}\right]\right) = \frac{1}{2}\left(\vec{r} \times \Delta\vec{r}\right) \quad . \tag{4.112}$$

The **surface velocity** then is:

$$\frac{d\vec{F}}{dt} = \frac{1}{2}(\vec{r} \times \vec{v}) = \frac{\vec{l}}{2\,m} = \text{const.} \tag{4.113}$$

(see Kepler's laws below).

4.3.7 External and Internal Angular Momentum

We introduce as coordinates:

- the centroid coordinate

$$\vec{r}_s = \frac{1}{M}\sum_{i=1}^{N} m_i\vec{r}_i \quad ; \quad M = \sum_{i=1}^{N} m_i \tag{4.114}$$

and

- the coordinates of the particles in the center of mass system

$$\vec{r}_i^{\ s} = \vec{r}_i - \vec{r}_s\,. \tag{4.115}$$

Then $\vec{L}$ can be rewritten as:

$$\vec{L} = \sum_i(\vec{r}_i^{\ s} + \vec{r}_s) \times (m_i\vec{v}_i^{\ s} + m_i\vec{v}_s) = \tag{4.116}$$

$$\sum_i(\vec{r}_i^{\ s} \times \vec{p}_i^{\ s}) + (\vec{r}_s \times \vec{p}_s) = \vec{L}_{\text{int}} + \vec{L}_s\ ,$$

using

$$\sum_i m_i\vec{r}_i^{\ s} = 0\,; \sum_i m_i\vec{v}_i^{\ s} = 0. \tag{4.117}$$

The first term in (4.116) $\vec{L}_{\text{int}}$ is called **internal angular momentum**; it is related to the center of mass system and independent of its motion in space, i.e. independent of the observer. The second term $\vec{L}_s$ is called **external angular**

momentum; it corresponds to the angular momentum of a particle of mass *M* and via $\overrightarrow{r_s}$ depends on the origin of the coordinate system, i.e. depends on the observer.

The change of $\overrightarrow{L}$ in time is:

$$\boxed{\frac{d\overrightarrow{L}}{dt} = \frac{d\overrightarrow{L}_{\text{int}}}{dt} + \frac{d\overrightarrow{L_s}}{dt} \ , \qquad (4.118)}$$

where

$$\frac{d\overrightarrow{L_s}}{dt} = \overrightarrow{r}_s \times \frac{d\overrightarrow{p_s}}{dt} = \overrightarrow{r}_s \times \overrightarrow{F}_a \ . \qquad (4.119)$$

If no external force is acting, $\overrightarrow{F}_a = 0$, then

$$\overrightarrow{L}_s = \text{const.} \ , \qquad (4.120)$$

and the change of $\overrightarrow{L}$ arises only from the change in $\overrightarrow{L}_{\text{int}}$.

To examine this change in more detail let's separate the torque as follows

$$\boxed{\overrightarrow{N} = \sum_i(\overrightarrow{r_i} \times [\overrightarrow{F}_{ia} + \sum_{j\neq i} \overrightarrow{F}_{ij}]) = \sum_i(\overrightarrow{r_i} \times \overrightarrow{F}_{ia}) + \sum_{i<j}(\overrightarrow{r_i} - \overrightarrow{r_j}) \times \overrightarrow{F}_{ij} \ . \qquad (4.121)}$$

Here $\overrightarrow{F}_{ia}$ is the external force acting on particle *i*, and it is used that $\overrightarrow{F}_{ij} = -\overrightarrow{F}_{ji}$ according to the actio=reactio principle. The further discussion will be carried out for the case that the internal forces are central forces, i.e. $\overrightarrow{F}_{ij}$ is parallel $\overrightarrow{r}_{ij} = \overrightarrow{r_i} - \overrightarrow{r_j}$. Then the 2nd term vanishes and we obtain:

$$\overrightarrow{N} = \sum_i(\overrightarrow{r_i} \times \overrightarrow{F}_{ia}) = \overrightarrow{N}_a \ ; \qquad (4.122)$$

i.e. the torque only arises from the external forces. For a **closed system** we have

$$\overrightarrow{F}_a = 0 \ ; \overrightarrow{N}_a = 0 \qquad (4.123)$$

such that

$$\overrightarrow{L} = \text{const} \ , \ \overrightarrow{L}_s = \text{const}, \text{also}\ \overrightarrow{L}_{\text{int}} = \text{const}. \qquad (4.124)$$

4.3.8 Exchange of Momentum and Angular Momentum in the Collision of Two (or Several) Particles

We consider the collision between two particles that interact by a central force; in this case there are no external forces. Angular momentum and momentum conservation give:

$$\overrightarrow{l_1} + \overrightarrow{l_2} = \overrightarrow{l_1}' + \overrightarrow{l_2}' \tag{4.125}$$

$$\text{before the collision} \qquad \text{after the collision}$$

$$\overrightarrow{p_1} + \overrightarrow{p_2} = \overrightarrow{p_1}' + \overrightarrow{p_2}' \, . \, .$$

For the change in momentum and angular momentum of particles 1 and 2 follows:

$$\Delta\overrightarrow{p_1} = -\Delta\overrightarrow{p_2} \quad : \text{momentum exchange} \tag{4.126}$$

and

$$\Delta\overrightarrow{l_1} = -\Delta\overrightarrow{l_2} \quad : \text{angular momentum exchange} \tag{4.127}$$

4.4 Energy

In addition to momentum and angular momentum, the **energy** provides essential information about a physical system. For many important cases energy is also a conserved quantity.

4.4.1 Kinetic Energy and Work

A mass point of mass m is moving on a trajectory $\overrightarrow{r}(t)$ under the influence of a force $\overrightarrow{F}$ from point a to b. The **work** done by the force $\overrightarrow{F}$ on the mass point along the trajectory from a to b (W_{ab}) is defined by the line integral

$$W_{ab} = \int_a^b \overrightarrow{F} \cdot d\overrightarrow{r}, \tag{4.128}$$

where the line integral is carried out along the particle trajectory $\overrightarrow{r}(t)$. Due to the scalar product the work is only determined by the component of the force in the direction of the path. The work W_{ab} then is a scalar quantity.

The connection between the work W_{ab} and the **kinetic energy** of the mass point follows from:

$$m\frac{d\overrightarrow{v}}{dt} = \overrightarrow{F} \quad . \tag{4.129}$$

By forming the scalar product with $\overrightarrow{v}$ and integrating in time we obtain:

$$\int_{t_a}^{t_b} m\left(\frac{d\overrightarrow{v}}{dt}\cdot\overrightarrow{v}\right)dt = \int_{t_a}^{t_b}\left(\overrightarrow{F}\cdot\overrightarrow{v}\right)\ dt\ . \tag{4.130}$$

The right side of this relationship is precisely the work:

$$\int_{t_a}^{t_b}\left(\overrightarrow{F}\cdot\overrightarrow{v}\right)\ dt = \int_{t_a}^{t_b} F_T v\ dt = \int_a^b F_T\ ds = \int_a^b \overrightarrow{F}\cdot d\overrightarrow{r}, \tag{4.131}$$

if the mass point is at point $a(b)$ at time $t_{a(b)}$. F_T denotes the component of the force $\overrightarrow{F}$ tangential to the trajectory and s is the arc length of the path. We can integrate the left hand side:

$$m\int_{t_a}^{t_b}\left(\frac{d\overrightarrow{v}}{dt}\cdot\overrightarrow{v}\right)dt = m\int_{t_a}^{t_b}\frac{d}{dt}\left(\frac{v^2}{2}\right)dt = \frac{m}{2}\left(v_b^2 - v_a^2\right) \tag{4.132}$$

with

$$v_a^2 = v(t_a)^2 \quad , \quad v_b^2 = v(t_b)^2 \quad . \tag{4.133}$$

Defining the **kinetic energy** T of a particle of mass m and of velocity $\overrightarrow{v}$ by:

$$\boxed{T = \frac{1}{2}mv^2 = \frac{p^2}{2m}\ , \tag{4.134}}$$

we find:

$$T_b - T_a = W_{ab}\ . \tag{4.135}$$

In words:

The work done by the force $\overrightarrow{F}$ along the trajectory $\overrightarrow{r}(t)$ from *a* to *b* is equal to the change in kinetic energy.

Example: Free fall
A body of mass *m* drops under the influence of constant gravity from the height z_0, where it is at rest at time $t = 0$ ($\overrightarrow{v}(0) = 0$). The work done by gravity then is:

$$W_{z_0 \to 0} = -\int_{z_0}^{0} mgdz = +mgz_0 \ ; \tag{4.136}$$

it is equal to the kinetic energy that is reached before impact with the earth's surface:

$$T_0 = \frac{m}{2} v_0^2 = mgz_0 \ , \tag{4.137}$$

since $T(0) = 0$ due to the initial condition.

Extension to a system of *N* particles:

The kinetic energy of a system of *N* particles is defined by

$$T = \sum_{i=1}^{N} T_i = \sum_{i=1}^{N} \frac{1}{2} m_i v_i^2 \ . \tag{4.138}$$

From the equations of motion

$$m_i \frac{d\overrightarrow{v_i}}{dt} = \overrightarrow{F_i} \tag{4.139}$$

we derive (as above):

$$T_b - T_a = \sum_i \int_{t_a}^{t_b} \overrightarrow{F_i} \cdot \overrightarrow{v_i} \ dt = \sum_i \int_a^b F_{Ti} \ ds_i = \sum_i W_{ab}^i = W_{ab} \, , \tag{4.140}$$

where *a* and *b* represent the position of the particles $\overrightarrow{r_i}$ at the times t_a and t_b. F_{Ti} is the component of the force $\overrightarrow{F_i}$ tangential to the path of the *i*-th particle; s_i the associated arc length.

4.4.2 Conservative Forces, Potential Energy, Energy Theorem

For the sake of simplicity we limit ourselves to a single mass point below. The definition of work generally depends not only on the integration boundaries *a*, *b*, but also on the shape of the trajectory (Fig. 4.2):

$$\underbrace{\int_{t_a}^{t_b} \vec{F} \cdot \vec{v}\, dt}_{\text{path 1}} \neq \underbrace{\int_{t_a}^{t_b} \vec{F} \cdot \vec{v}\, dt}_{\text{path 2}} \tag{4.141}$$

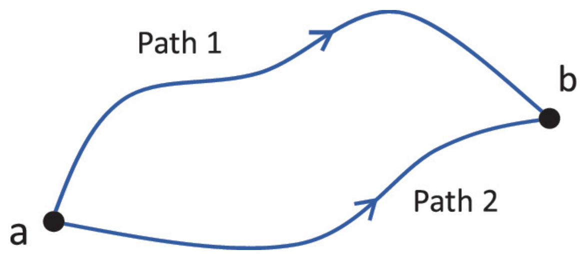

Fig. 4.2 Illustration of two different trajectories connecting the points *a* and *b*

Particular forces have proven to be important in physics, where W_{ab} is independent of the trajectory between *a* and *b*. We refer to such forces as **conservative forces**. In the mathematical sense we call a force **conservative**, if a scalar function $U(\vec{r})$ exists such that:

$$W_{ab} = \int_a^b \vec{F} \cdot \vec{ds} = U(a) - U(b)\ . \tag{4.142}$$

The function $U(\vec{r})$ is called **potential energy** of the particle at position $\vec{r}$. It is only determined up to an additive constant.

Conclusions:

Work along a closed path

For a conservative force the integration over any closed path gives:

$$\oint \vec{F} \cdot \vec{ds} = \oint F_T\ ds = 0\ . \tag{4.143}$$

Energy theorem

For a conservative force we get:

$$T_b + U(b) = T_a + U(a)\ . \tag{4.144}$$

The total energy of the particle

$$E = T + U \tag{4.145}$$

therefore is constant.

Example: mass point under the influence of gravity

$$\int_{t_a}^{t_b} \vec{F} \cdot \vec{v} \, dt = -\int_a^b mg \; dz = mgz_a - mgz_b = U(a) - U(b)\,. \tag{4.146}$$

Since U is only fixed up to an additive constant we can fix U by the condition that at the earth's surface $U(0) = 0$. Then the potential energy $U(h) = mgh$ of the particle at the height h above the earth's surface is equal to the work done (against gravity) to raise the mass point from the earth's surface to the height h without changing its kinetic energy. If the mass point drops from a height h **free** we get

$$E = \frac{1}{2}mv^2 + mgz = \text{const.} = mgh \tag{4.147}$$

for every point on the trajectory, if the mass point was at rest at $z = h$. **The increase in kinetic energy is equal to the decrease in potential energy.**

Calculation of the force $\vec{F}$ from the potential energy $U(\vec{r})$

$$\vec{F} = -\begin{pmatrix} \frac{\partial U}{\partial x} \\ \frac{\partial U}{\partial y} \\ \frac{\partial U}{\partial z} \end{pmatrix} = -\operatorname{grad} U = -\vec{\nabla} U\,. \tag{4.148}$$

Here $\partial U/\partial x$ is the **partial derivative** of the function $U = U(x, y, z)$ for fixed values of y, z.

Proof

1. From Eq. (4.148) we get:

$$W_{ab} = \int_a^b F_T \; ds = U(a) - U(b)\,. \tag{4.149}$$

since:

$$\int_{t_a}^{t_b} \vec{F} \cdot \vec{v} \, dt = -\int_a^b (\vec{\nabla} U \cdot \vec{dr}) =$$

$$-\int_a^b \left(\frac{\partial U}{\partial x} dx + \frac{\partial U}{\partial y} dy + \frac{\partial U}{\partial z} dz \right) = -\int_a^b dU = U(a) - U(b)\,. \tag{4.150}$$

The **total differential** dU is the change of U in the transition from the point $\overrightarrow{r}$ to the infinitesimally neighboring point $\overrightarrow{r} + \overrightarrow{dr}$:

$$dU = \frac{\partial U}{\partial x} dx + \frac{\partial U}{\partial y} dy + \frac{\partial U}{\partial z} dz \ = \mathrm{grad\,U} \cdot \overrightarrow{\mathrm{dr}}. \tag{4.151}$$

2. If a function U exists which fulfills

$$\boxed{W_{ab} = \int_a^b F_T \ ds = U(a) - U(b),} \tag{4.152}$$

the total energy of the particle is

$$\boxed{E = \frac{p^2}{2\,m} + U(\overrightarrow{r}).} \tag{4.153}$$

The conservation of energy follows from the time derivative of E:

$$\frac{dE}{dt} = \frac{d}{dt}\left(\frac{p^2}{2\,m}\right) + \frac{d}{dt} U(x, y, z) =$$

$$v_x\left(\frac{dp_x}{dt} + \frac{\partial U}{\partial x}\right) + v_y\left(\frac{dp_y}{dt} + \frac{\partial U}{\partial y}\right) + v_z\left(\frac{dp_z}{dt} + \frac{\partial U}{\partial z}\right) = 0\,. \tag{4.154}$$

If the force $\overrightarrow{F}$ does not depend on the velocity, the ( ) brackets are independent of $\overrightarrow{v}$. Since $\overrightarrow{v}$ may have arbitrary values it follows that

$$\frac{dp_x}{dt} = F_x = -\frac{\partial U}{\partial x} \text{ etc. for } y, z\,. \tag{4.155}$$

For a conservative system of N particles we get:

$$\boxed{\overrightarrow{F_i} \ = \ -\,\mathrm{grad_i\,U} = -\overrightarrow{\nabla}_{\mathrm{i}}\mathrm{U}} \tag{4.156}$$

for the force acting on particle i, with

$$U = U(\overrightarrow{r}_1, \overrightarrow{r}_2, \ldots \overrightarrow{r}_N)\ . \tag{4.157}$$

Example:

The potential energy of a particle is given by:

$$U = \frac{a}{r} + b \; , \tag{4.158}$$

with

$$r^2 = x^2 + y^2 + z^2 \; . \tag{4.159}$$

The associated force is a central force:

$$\overrightarrow{F} = -\,\mathrm{grad}\left(\frac{a}{r}\right) = a\frac{\overrightarrow{r}}{r^3} \tag{4.160}$$

using

$$\frac{\partial}{\partial x}\left(\frac{1}{r}\right) = \frac{d}{dr}\left(\frac{1}{r}\right)\cdot\frac{\partial r}{\partial x} = \left(\frac{-1}{r^2}\right)\frac{x}{r}, \tag{4.161}$$

$$\frac{\partial}{\partial y}\left(\frac{1}{r}\right) = \frac{d}{dr}\left(\frac{1}{r}\right)\cdot\frac{\partial r}{\partial y} = \left(\frac{-1}{r^2}\right)\frac{y}{r},$$

$$\frac{\partial}{\partial z}\left(\frac{1}{r}\right) = \frac{d}{dr}\left(\frac{1}{r}\right)\cdot\frac{\partial r}{\partial z} = \left(\frac{-1}{r^2}\right)\frac{z}{r}.$$

4.4.3 Invariances of *U*; Separation of Center of Mass Energy

Translational invariance

The property

$$U(\overrightarrow{r_i}) = U(\overrightarrow{r_i} + \overrightarrow{a}) \tag{4.162}$$

for any vectors $\overrightarrow{a}$ implies that U may only depend on the internal coordinates of the system (of N particles), e.g. on the distance vectors

$$\overrightarrow{r}_{ij} = \overrightarrow{r}_i - \overrightarrow{r}_j \; , \tag{4.163}$$

such that

$$U = U(\overrightarrow{r_{ij}}) \; . \tag{4.164}$$

From $\overrightarrow{\nabla}_i U = -\overrightarrow{\nabla}_j U$ we obtain

(4.165)

$$\overrightarrow{F}_{ij} = -\overrightarrow{F}_{ji}$$

since $\overrightarrow{r}_{ij} = -\overrightarrow{r}_{ji}$. This is the **actio=reactio principle**, from which, together with the equations of motion, we have derived the momentum conservation law. The **momentum conservation** is therefore a direct consequence of the translational invariance of *U*.

Rotational invariance
In case of rotational invariance we have

$$U(\overrightarrow{r_i}) = U(\overrightarrow{r_i}\,') \ , \tag{4.166}$$

where $\overrightarrow{r_i}\,'$ emerges from $\overrightarrow{r_i}$ by an arbitrary rotation. It follows that *U* is only a function of the distances

$$r_{ij} = |\overrightarrow{r_i} - \overrightarrow{r_j}| \tag{4.167}$$

i.e.,

$$U = U(r_{ij}) \ . \tag{4.168}$$

The force acting between 2 particles *i*, *j* then is a central force:

$$\overrightarrow{F}_{ij} = \tilde{f}(r_{ij})\overrightarrow{r}_{ij} \ , \tag{4.169}$$

since for any function *f*(*r*) holds:

$$\frac{\partial}{\partial x} f(r) = \frac{df}{dr}\frac{\partial r}{\partial x} = \frac{df}{dr}\frac{x}{r} = g(r)x \ ; \quad \text{also for} \ \ y, z \ . \tag{4.170}$$

The **angular momentum conservation** holds for central forces, which is therefore a consequence of the rotational invariance.

Invariance with respect to time translations
When discussing the conservation of energy, we have assumed that *U* does not depend explicitly on time *t*,

$$\frac{\partial U}{\partial t} = 0 \ . \tag{4.171}$$

This equation can also be understood as a consequence of the invariance of *U* against time translations, $t \to t + \Delta t$ for any Δt. The **energy conservation law** is therefore a consequence of the invariance with respect to time translations.

Galilean invariance

In this case the scalar function $U = U(\overrightarrow{r_{ij}})$ does not change for a Galilean transformation. The kinetic energy is:

$$T' = \frac{1}{2}\sum_i m_i(\overrightarrow{v_i} - \overrightarrow{v})^2 = T - \overrightarrow{P}\cdot\overrightarrow{v} + \frac{1}{2}Mv^2 \ . \tag{4.172}$$

Since for a system with $U = U(\overrightarrow{r_{ij}})$ the momentum $\overrightarrow{P} = \sum_i m_i \cdot \overrightarrow{v_i}$ is conserved, the kinetic energy only changes by an additive constant,

$$T' = T + \text{const} \ , \tag{4.173}$$

i.e. the **energy** for a closed system

$$\boxed{E = T + U = \text{const}} \tag{4.174}$$

is **Galilean-invariant** like the momentum and angular momentum.

If we specifically choose the coordinate system Σ as the center of mass system, $\overrightarrow{P} = 0$, we get:

$$\boxed{T' = T + \frac{1}{2}Mv^2 = T_{\text{int}} + T_s \ .} \tag{4.175}$$

T_{int} here is the internal kinetic energy, T_s the center of mass energy with respect to the system Σ' with velocities $\overrightarrow{v_i}\,'$. Since $U = U(\overrightarrow{r_{ij}})$ in the transition $\Sigma \rightarrow \Sigma'$ does not change, we can always separate the center of mass energy in a closed system,

$$\boxed{E = T_s + E_{\text{int}} \ ,} \tag{4.176}$$

where E_{int} is the energy in the center of mass system.

4.4.4 Friction Forces

All **fundamental** forces known to us are conservative in the sense of equation

$$W_{ab} = \int_a^b F_T \ ds = U(a) - U(b) \ , \tag{4.177}$$

i.e. the energy law applies. This includes the case of the **Lorentz force** (force of a magnetic field $\overrightarrow{B}$ on a charge q moving with velocity $\overrightarrow{v}$),

$$\overrightarrow{F} = \frac{q}{c}(\overrightarrow{v} \times \overrightarrow{B}) \ . \tag{4.178}$$

Since $\overrightarrow{F}$ is always perpendicular to the direction of motion,

$$\overrightarrow{F} \cdot \overrightarrow{v} = \frac{q}{c}(\overrightarrow{v} \times \overrightarrow{B}) \cdot \overrightarrow{v} = 0 \ , \tag{4.179}$$

it doesn't do any work; thus it doesn't affect the energy balance.

Additionally, there are forces that enter into the energy balance and increase the loss of energy in the system: **frictional forces**. They become important for the description of the motion of a body in a gas or a liquid or on a surface (sliding friction). In the simplest case frictional forces are proportional to $\overrightarrow{v}$:

$$\boxed{\overrightarrow{F_r} = -c\overrightarrow{v} \, ; \, c > 0 \ .} \tag{4.180}$$

Then the system suffers a loss of energy because of

$$\boxed{\int_{t_a}^{t_b} \overrightarrow{F}_R \cdot \overrightarrow{v} \, dt = -c \int_{t_a}^{t_b} v^2 \ \ dt < 0.} \tag{4.181}$$

The occurrence of frictional forces does not contradict the statement that all fundamental forces are conservative because frictional forces are not conservative forces, but a result from a more general description of the interaction, e.g. between the molecules of a rolling ball and those of the surface where the ball is rolling.

Addition: vector property of *gradU*

1.
Addition

If $U(\overrightarrow{r}) = U_1(\overrightarrow{r}) + U_2(\overrightarrow{r})$ it follows from the rules of differentiation:

$$\boxed{\operatorname{grad} U = \operatorname{grad} U_1 + \operatorname{grad} U_2 \ ,} \tag{4.182}$$

the vector addition law defined for vectors. It also holds for multiplication by a real number α,

$$\alpha \ \operatorname{grad} U = \operatorname{grad}(\alpha U) \ . \tag{4.183}$$

2. **Transformation behavior for rotations**

   The scalar function $U(\vec{r})$ assigns a real number to each point in space $\vec{r}$ that does not change when the coordinate system is rotated. Accordingly for a scalar function U under rotations we have:

$$U(x_1, x_2, x_3) = U'(x_1', x_2', x_3') \ , \tag{4.184}$$

where the components of $\vec{r}$ (see Sect. 2.3.2) are rotated with the matrix d_{ij} like:

$$x_i' = \sum_j d_{ij} x_j \quad \text{with} \quad \sum_i d_{im} d_{in} = \delta_{mn} \ . \tag{4.185}$$

It follows from the chain rule for differentiation:

$$\frac{\partial U'(x_1', x_2', x_3')}{\partial x_i'} = \sum_j \frac{\partial U(x_1, x_2, x_3)}{\partial x_j} \frac{\partial x_j}{\partial x_i'} = \sum_j d_{ij} \frac{\partial U(x_1, x_2, x_3)}{\partial x_j} \ , \tag{4.186}$$

i.e. the components of grad U transform under rotations like the components of $\vec{r}$. In Eq. (4.186) we have used:

$$\sum_i d_{ik} x_i' = \sum_{i,j} d_{ik} d_{ij} x_j = \sum_j \delta_{kj} x_j = x_k \tag{4.187}$$

employing (4.185).

In summarizing this chapter we have defined forces and derived Newton’s equations of motion; their solution provides the trajectory of a mass point in space and time. Examples for characteristic problems have been given and the explicit solutions been derived in detail. We have found that instead of velocities or angular velocities it is more convenient to introduce momenta and angular momenta of particles since for closed systems - without external forces—the total momentum is a constant of motion. This also holds for the angular momentum if no external torque acts on the system and is a consequence of Galilean invariance. We have examined the connection between the work done by a force on a particle along its trajectory and the actual kinetic energy. In case of conservative forces one can introduce a potential energy $U(\overrightarrow{r})$ that allows to compute the actual force by its negative gradient. Then the energy of the system is defined by the sum of kinetic and potential energy and—for closed systems—is found to be a conserved quantity, too. The energy balance, however, does not hold for frictional forces which are some 'effective' forces that result from a more general description of the microscopic interactions, e.g. between the molecules of a rolling ball and those of the surface.

W. Cassing, *Theoretical Physics Compact I*
https://doi.org/10.1007/978-3-031-95512-9_5

5. Applications of Newton Mechanics

Wolfgang Cassing[1]
(1) University of Giessen, Gießen, Hessen, Germany

In this chapter we will continue with applications of Newtonian mechanics for central forces, where the potential U only depends on the magnitude of the relative distance $|\overrightarrow{r}_1 - \overrightarrow{r}_2|$ between two mass points. In this case the conservation of momentum, angular momentum and energy holds, which drastically reduces the number of free degrees of freedom. An important case are $1/r^2$-forces, which holds for Coulomb and gravitational forces; we will classify the trajectories according to their energy and derive Kepler's laws for the motion of planets. In extension the law of gravity will be derived and gravity fields are introduced for static mass distributions. In addition the dynamics of a linear oscillator is discussed—another important physical system—and the solutions are calculated from the equations of motion also in case of additional frictional forces. The case of a damped oscillator, that is driven by an external periodic force, will lead to the formation of resonances that are analysed in some detail. In addition the problem of coupled harmonic oscillations is addressed, which is characteristic for the vibrational modes in crystals.

5.1 Central Forces

One of the most important problems in theoretical physics is the motion of 2 mass points under the influence of a central force with applications in astrophysics, atomic physics and nuclear physics.

5.1.1 Reduction of Degrees of Freedom

We consider a closed system of two particles without any external forces,

$$\overrightarrow{F}_a = 0 \ . \tag{5.1}$$

A central force acts between the particles

$$\overrightarrow{F}_{12} = -\operatorname{grad} U = f(r)\frac{\overrightarrow{r}}{r} = -\overrightarrow{F}_{21} \tag{5.2}$$

with

$$\overrightarrow{r} = \overrightarrow{r}_{12} = \overrightarrow{r}_1 - \overrightarrow{r}_2 = -\overrightarrow{r}_{21}. \tag{5.3}$$

The equations of motion in the coordinates $\overrightarrow{r}_1, \overrightarrow{r}_2$,

$$m_1\overrightarrow{a}_1 = \overrightarrow{F}_{12}, \qquad m_2\overrightarrow{a}_2 = \overrightarrow{F}_{21} \tag{5.4}$$

can be rewritten in center of mass and relative coordinates ($M = m_1 + m_2$):

$$\overrightarrow{r}_s = \frac{1}{M}(m_1\overrightarrow{r}_1 + m_2\overrightarrow{r}_2), \qquad \overrightarrow{r} = \overrightarrow{r}_1 - \overrightarrow{r}_2. \tag{5.5}$$

From

$$m_1\overrightarrow{a}_1 + m_2\overrightarrow{a}_2 = \overrightarrow{F}_{12} + \overrightarrow{F}_{21} = 0 \tag{5.6}$$

we get

$$\frac{d^2}{dt^2}\overrightarrow{r}_s = \overrightarrow{a}_s = 0. \tag{5.7}$$

The solution is known: this is a straight, uniform motion for the center of mass.

For the relative motion one obtains (by taking the difference)

$$\overrightarrow{a}_1 - \overrightarrow{a}_2 = \frac{\overrightarrow{F}_{12}}{m_1} - \frac{\overrightarrow{F}_{21}}{m_2} = \left(\frac{1}{m_1} + \frac{1}{m_2}\right)\overrightarrow{F}_{12}, \tag{5.8}$$

or

$$\mu\ddot{\overrightarrow{r}} = \overrightarrow{F}_{12} = \overrightarrow{F} \tag{5.9}$$

with the **reduced mass** μ,

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2} = \frac{m_1+m_2}{m_1 m_2}. \tag{5.10}$$

This reduces the two-body problem to the **equivalent** one-body problem for a fictitious particle of mass μ under the influence of the force $\vec{F}$. Instead of 6 differential equations there are only 3 differential equations to solve.

With the help of the energy and angular momentum theorem, the problem can be reduced to only a single degree of freedom (in the variable *r*). From the conservation of angular momentum

$$\vec{l} = \text{const} \tag{5.11}$$

it follows that the motion proceeds in a plane. Thus we can - without restrictions - employ the parameter representation (in the *x*, *y* plane)

$$\vec{r} = \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \\ 0 \end{pmatrix}. \tag{5.12}$$

Furthermore, only the energy of the internal motion is of interest,

$$E_{int} = \frac{1}{2}\mu v^2 + U(r), \tag{5.13}$$

which we can rewrite as:

$$E_{int} = \frac{1}{2}\mu\dot{r}^2 + \frac{l^2}{2\mu r^2} + U(r). \tag{5.14}$$

This equation only contains a single variable (*r*) and its time derivative ($\dot{r}$) (for fixed *l*).

Proof For the velocity we get from (5.12)

$$\vec{v} = \begin{pmatrix} \dot{r}\cos\varphi \\ \dot{r}\sin\varphi \\ 0 \end{pmatrix} + \begin{pmatrix} -r\dot{\varphi}\sin\varphi \\ r\dot{\varphi}\cos\varphi \\ 0 \end{pmatrix} = \dot{r}\vec{e}_r + r\dot{\varphi}\vec{e}_\varphi. \tag{5.15}$$

Since

$$\tag{5.16}$$

$$\overrightarrow{e_r} \cdot \overrightarrow{e_\varphi} = 0,$$

we obtain

$$E = \frac{\mu}{2}\left(\dot{r}\overrightarrow{e_r} + r\dot{\varphi}\overrightarrow{e_\phi}\right)^2 + U(r) = \frac{\mu}{2}(\dot{r}^2 + r^2\dot{\varphi}^2) + U(r). \tag{5.17}$$

The angle variable $\dot{\varphi}$ can be eliminated using the magnitude of the angular momentum $\overrightarrow{l}$, which is constant in time:

$$l = \mu|\overrightarrow{r} \times \overrightarrow{v}| = \mu r^2 \dot{\varphi} \qquad \text{q. e. d.} \tag{5.18}$$

Note: The total angular momentum of the two particles can be decomposed in an 'outer' (center of mass) part and an 'inner' part. For central forces both parts are separately conserved in the absence of external forces. $\overrightarrow{l}$ denotes the internal part, i.e. the relative angular momentum of the two particles. Equation (5.17) can be interpreted as energy for a 1-dimensional motion in the variable r with an **effective** potential energy

$$\boxed{U^l_{\text{eff}} = \frac{l^2}{2\mu r^2} + U(r),} \tag{5.19}$$

thus

$$\boxed{E = \frac{1}{2}\mu\dot{r}^2 + U^l_{\text{eff}}(r).} \tag{5.20}$$

The term from the kinetic energy $l^2/(2\mu r^2) = U_z$ is called **centrifugal potential** and is added to the potential energy.

To explain the term "centrifugal potential" we calculate the associated force,

$$\overrightarrow{F}_z = -\operatorname{grad} U_z = \frac{l^2}{\mu r^3}\overrightarrow{e_r} = \mu r \omega^2 \overrightarrow{e_r}, \tag{5.21}$$

i.e. the product of μ and the centrifugal acceleration.

5.1.2 Classification of Trajectories

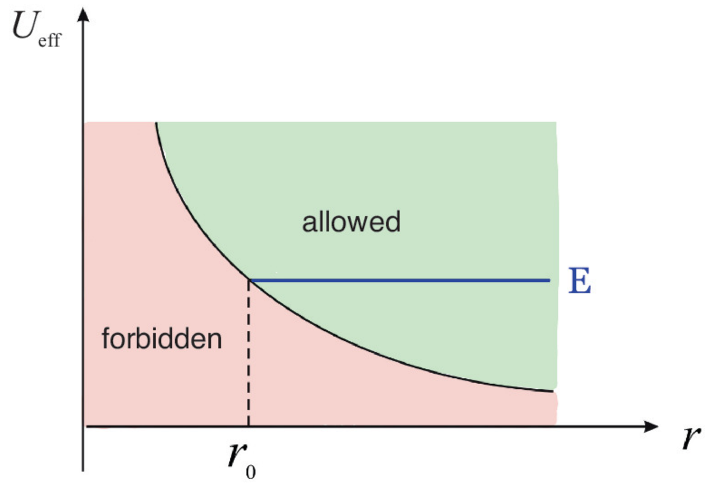

Fig. 5.1 Example for a potential U_{eff} that is positive everywhere and decreasing with r

$dU/dr < 0$ **for all** r

Since U_z is also repulsive everywhere, $U^l_{\text{eff}}(r)$ (with the convention $U^l_{\text{eff}}(\infty) = 0$) has the following qualitative shape (Fig. 5.1):

For fixed energy E only orbits with $r \geq r_0$ are possible, since for $r < r_0$ the kinetic energy T_r is negative, i.e. the velocity $\dot{r}$ will be imaginary. The permitted trajectories are called **unbound states** or **scattering states**.

$d^2U/dr^2 > 0$ **for all** r

1. $\lim_{r\to\infty} U_{\text{eff}}(r) \to \infty$ (Fig. 5.2).

   Since $T_r > 0$ must always be positive we only get **bound states** for $r_1 \leq r \leq r_2$.

2. Normalization: $U_{\text{eff}}(\infty) = 0$ (Fig. 5.3).

   For $E > 0$ one gets unbound states (green dotted area), bound states for $E < 0$ (green area).

3. Normalization: $U_{\text{eff}}(\infty) = 0$ (Fig. 5.4).

   For $E > U_m$ there are only unbound states for arbitrary $r \geq 0$. If $0 \leq E < U_m$ both bound and unbound states exist. For $E < 0$ there are only bound states.

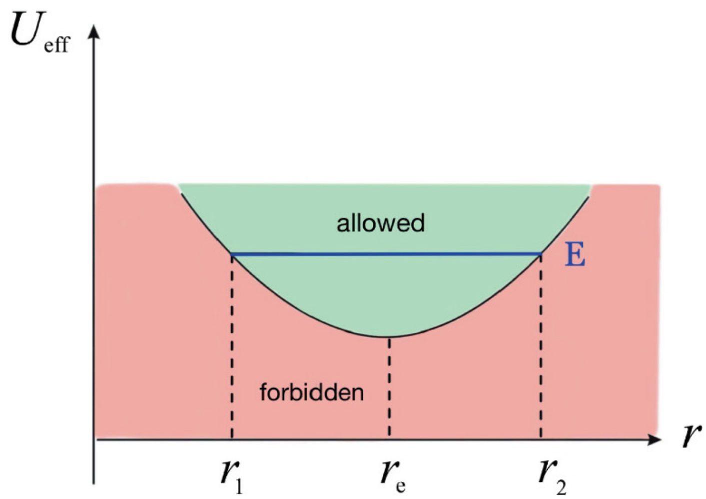

Fig. 5.2 Example for a potential that only allows for bound states

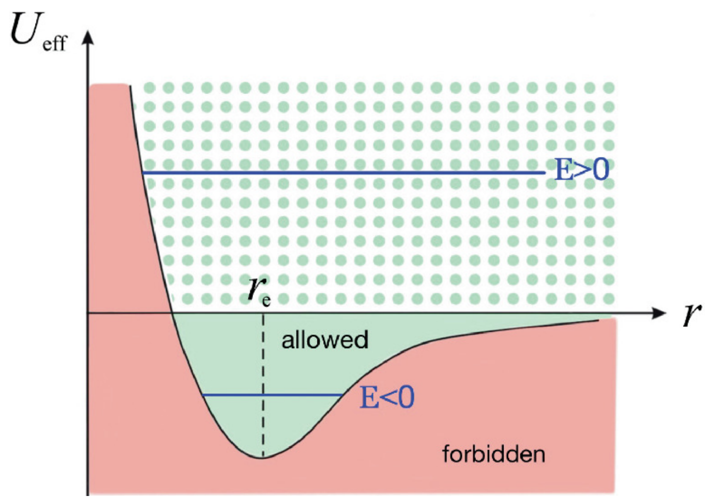

Fig. 5.3 Example for potential that allows for bound states ($E < 0$) as well as scattering states ($E > 0$)

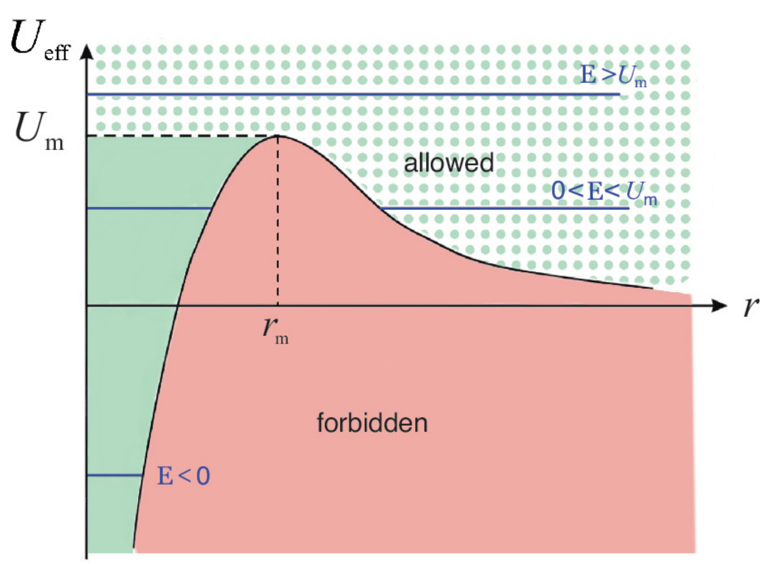

Fig. 5.4 Example for a potential with only unbound states for arbitrary $r \geq 0$ and $E > U_m$ (green dots). For $0 \leq E < U_m$ both bound (green) and unbound states (green dots) exist, while for $E < 0$ only bound states can appear

Equilibrium: In cases (1.) and (2.) there is no force for $r = r_e$ since

$$\left(\frac{dU_{\text{eff}}}{dr}\right)_{r=r_e} = 0. \tag{5.22}$$

The same applies to case 3.) at $r = r_m$. In these points the system is in equilibrium.

In cases (1.) and (2.) this equilibrium is **stable** but in case (3.) the equilibrium is **unstable**: for a small deflection from the equilibrium position a repulsive force acts, which drives the particle away from the equilibrium position at r_m.

5.1.3 $1/r^2$–Forces

For the practically important case

$$U = \pm\frac{c}{r}; \qquad c > 0 \tag{5.23}$$

we want to determine the trajectories explicitly.

The internal energy in this case is

$$\boxed{E = \frac{1}{2}\mu\dot{r}^2 + \frac{l^2}{2\mu r^2} \pm \frac{c}{r} = \text{const}} \tag{5.24}$$

and energy conservation gives:

$$\frac{dE}{dt} = 0 = \dot{r}\left(\mu\ddot{r} - \frac{l^2}{\mu r^3} \mp \frac{c}{r^2}\right). \tag{5.25}$$

Since in general $\dot{r} \neq 0$, the equation of motion follows as:

$$\boxed{\mu\ddot{r} - \frac{l^2}{\mu r^3} \mp \frac{c}{r^2} = 0.} \tag{5.26}$$

For $l = 0$ the motion occurs along a straight line ($\dot{\varphi} = 0$):

$$\vec{l} = 0 \longrightarrow \vec{r} \| \vec{v}. \tag{5.27}$$

To find the possible trajectories $r = r(\varphi)$ for $l \neq 0$, we introduce the new variable

$$w = \frac{1}{r} \text{with } \frac{dw}{d\varphi} = \frac{dw}{dr}\frac{dr}{d\varphi} = -\frac{1}{r^2}\frac{dr}{d\varphi} \tag{5.28}$$

and with ($\dot{\varphi} = l/(\mu r^2)$) we find

$$\dot{r} = \frac{dr}{d\varphi}\frac{d\varphi}{dt} = \dot{\varphi}\frac{dr}{d\varphi} = \frac{l}{\mu r^2}\frac{dr}{d\varphi} = -\frac{l}{\mu}\frac{dw}{d\varphi} \tag{5.29}$$

as well as

$$\ddot{r} = -\frac{l}{\mu}\frac{d}{dt}\frac{dw}{d\varphi} = -\frac{l}{\mu}\frac{d^2w}{d\varphi^2}\dot{\varphi} = -\frac{l^2}{\mu^2 r^2}\frac{d^2w}{d\varphi^2}. \tag{5.30}$$

Then the equation of motion (5.26) turns into:

$$-\frac{l^2}{\mu r^2}\left(\frac{d^2w}{d\varphi^2} + w \pm \frac{\mu c}{l^2}\right) = 0 \tag{5.31}$$

or ($l^2 \neq 0$)

$$\frac{d^2w}{d\varphi^2} + w = \mp\frac{\mu c}{l^2}. \tag{5.32}$$

The solution of the inhomogeneous differential equation of 2nd order (5.32) is the sum of the general solution of the homogeneous differential equation for $\tilde{w}$,

$$\frac{d^2\tilde{w}}{d\varphi^2} + \tilde{w} = 0, \tag{5.33}$$

given by

$$\tilde{w} = A \cos\varphi + B \sin\varphi = a \cos(\varphi - \varphi_0), \tag{5.34}$$

and an arbitrary solution of the inhomogeneous equation. A particular solution (for $d^2w/d\varphi^2 = 0$) is

$$w = \mp\frac{\mu c}{l^2}. \tag{5.35}$$

The general solution of (5.32) then reads:

$$\boxed{w = a \cos(\varphi - \varphi_0) \mp \frac{\mu c}{l^2},} \tag{5.36}$$

or with (5.28)

$$r(\mp 1 + \varepsilon \cos(\varphi - \varphi_0)) = p, \tag{5.37}$$

using the abbreviations

$$\varepsilon = \frac{al^2}{\mu c}, \qquad p = \frac{l^2}{\mu c}. \tag{5.38}$$

The integration constant a or ε is determined by the energy. Using elementary algebra we obtain (5.37)

$$\dot{r} = \dot{\varphi}\frac{dr}{d\varphi} = \frac{l}{\mu r^2}\frac{p\ \varepsilon \sin(\varphi-\varphi_0)}{(\mp 1+\varepsilon\cos(\varphi-\varphi_0))^2} \tag{5.39}$$

and (after a somewhat lengthy calculation) for the energy

$$E = \frac{\mu}{2}\dot{r}^2 + \frac{l^2}{2\mu r^2} \pm \frac{c}{r} = \frac{\mu c^2}{2l^2}(\varepsilon^2 - 1). \tag{5.40}$$

Equation (5.37) is the general form of a **conic section**. By appropriately choosing the coordinate system to which (r, φ) refers, we can rewrite (5.37) in the normal form

$$r(\mp 1 + \varepsilon \cos \varphi) = p, \quad \varepsilon \geq 0. \tag{5.41}$$

We consider different cases:

1. $U = -c/r$: **attraction**, i.e. $r(\varphi) = p/(1 + \varepsilon \cos \varphi)$. Then the following cases are possible:
   (a) $\varepsilon = 0$: circle; there is a bound state with $E < 0$.
   (b) $0 < \varepsilon < 1$: ellipse; there is also a bound state with $E < 0$ (Fig. 5.5).
   (c) $\varepsilon = 1$: parabola; in this case $E = 0$, it is an unbound state.

(d)
$\varepsilon > 1$: branch of a hyperbola that encloses the origin $r = 0$; unbound state with $E > 0$ (Fig. 5.6).

2.
$U = c/r$: **repulsion**, i.e. $r(\varphi) = p/(-1 + \varepsilon \cos \varphi)$.

We must require $\varepsilon > 1$, otherwise r would turn negative, and get the branch of the hyperbola that is complementary to the case 1.d) (Fig. 5.7).

Examples:

- Atomic systems:
  An example for case (2.) is electron-electron or proton-proton scattering. For the electron-proton system the orbits from case (1.) are possible, i.e. there can be bound states as well as scattering states (depending on the energy E).
- Planetary motion (see Sect. 5.2).

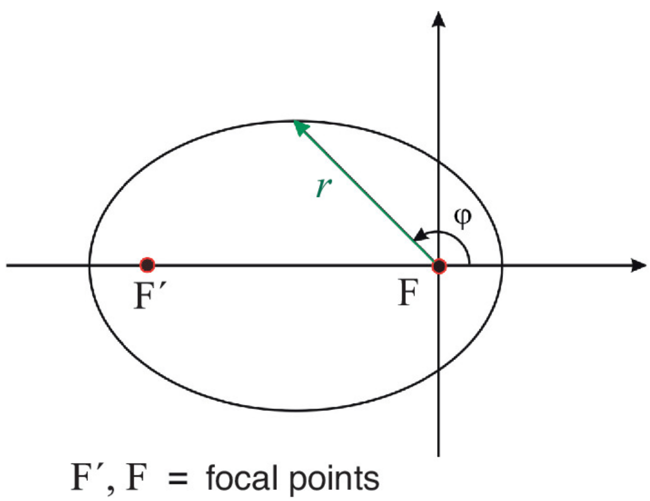

Fig. 5.5 Case of an ellipse, which is a bound state with $E < 0$. The center of mass is located in the focal point F

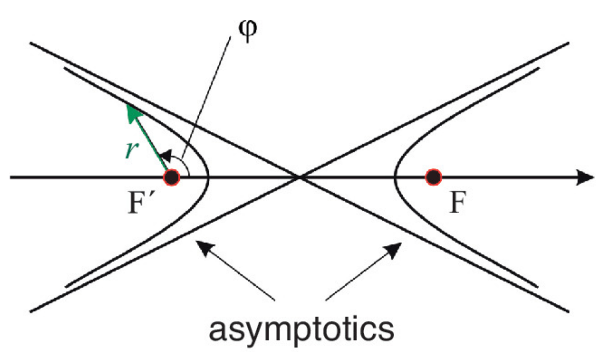

Fig. 5.6 Branch of a hyperbola that encloses the origin $r = 0$ and shows an unbound state with $E > 0$. The center of mass is located in the focal point F'

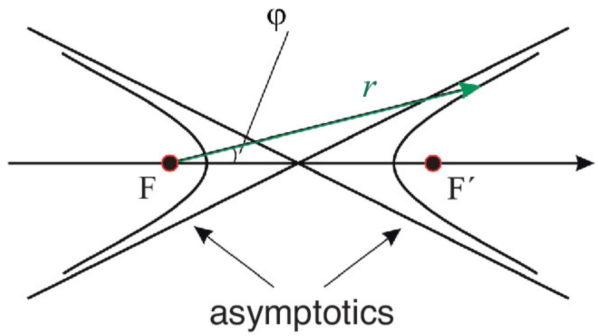

Fig. 5.7 Complementary branch of a hyperbola that does not include the origin $r = 0$ and shows an unbound state with $E > 0$. The center of mass is located in the focal point F

5.2 Planetary Motion; Gravity

5.2.1 Kepler's Laws

Kepler's laws describe the kinematics of the motion of planets:

1. **The planetary orbits are ellipses with the sun in a focal point.**
2. **The radius vector from the sun to the planet passes the same areas in equal times.**
3. **The squares of the orbital periods of different planets behave like the cubes of the semi-major axis of their elliptical orbits.**

The second law is the **area law** and shows, together with the 1st law containing the statement that the orbits are planar, that the angular momentum is conserved. The force responsible for planetary motion is a central force. Since the orbits are ellipses with the center of force in one of the focal points, we conclude from Sect. 4.4.2 that the central force is of the form

$$\overrightarrow{F} = -\tfrac{c}{r^2}\tfrac{\overrightarrow{r}}{r} = -\tfrac{c}{r^2}\overrightarrow{e}_r, \tag{5.42}$$

i.e. the potential energy is of the form

$$U(r) = -\tfrac{c}{r}, \qquad c > 0. \tag{5.43}$$

These equations therefore are the dynamical basis for Kepler's laws (1.) and (2.).

To explain the third law we use the area law

$$\frac{dF}{dt} = \frac{l}{2\mu}, \tag{5.44}$$

where we replaced the mass *m* by the reduced mass μ. Integration in time gives:

$$F = \frac{l}{2\mu}T, \tag{5.45}$$

where *T* is the orbital period and *F* is the area of the ellipse:

$$F = \pi ab \to T = \frac{2\pi\mu}{l}ab, \tag{5.46}$$

if *a* is the major axis and *b* the minor axis of the ellipse. Replacing

1. $r' + r = 2a$ (definition of the ellipse)
2. $a^2 = b^2 + c^2 = b^2 + \varepsilon^2 a^2$ with $\varepsilon = c/a$ (Pythagoras)
3. $(2a - r)^2 = r'^2 = r^2 + 4c^2 + 4cr \cos\varphi$ (cosine theorem according to 1.)
4. $r(1 + \varepsilon \cos\varphi) = (a^2 - c^2)/a = b^2/a = p$

and using

$$l^2 = \mu cp = \mu c\frac{b^2}{a}, \tag{5.47}$$

we obtain:

$$\boxed{T^2 = \frac{4\pi^2 a^2 b^2 \mu^2}{l^2} = \frac{4\pi^2\mu}{c}a^3.} \tag{5.48}$$

According to Kepler the pre-factor should be $4\pi^2\mu/c$ and be equal for all planets. To check this, let's look at the general

5.2.2 Law of Gravity

after which any 2 (electrically neutral) mass points at a distance *r* attract each other by a central force

$$\tag{5.49}$$

$$\vec{F} = -\frac{\gamma_1 \gamma_2}{r^2} \frac{\vec{r}}{r}.$$

Here γ_1 and γ_2 are characteristic constants for the mass points that are proportional to the masses m_1 and m_2 (in the equation of motion). This statement is by no means trivial, but follows from experiment, e.g. the **free fall**: For a freely falling body (near the earth's surface) we find

$$ma = -\frac{\gamma \gamma_E}{R_E^2}, \tag{5.50}$$

where *m* is the **inertial mass** of the body, γ and γ_E the constants for the body and the earth; R_E is the radius of the earth. If we now compare the free fall of two bodies 1 and 2, we have:

$$\frac{m_1 a_1}{m_2 a_2} = \frac{\gamma_1}{\gamma_2}. \tag{5.51}$$

Since one always finds $a_1 = a_2$ experimentally, we obtain

$$\frac{m_1}{m_2} = \frac{\gamma_1}{\gamma_2}. \tag{5.52}$$

Accordingly the mass *m* and the factor γ differ by only a universal constant factor such that the force can also be written as:

$$\vec{F} = -\Gamma \frac{m_1 m_2}{r^2} \frac{\vec{r}}{r} \tag{5.53}$$

for two bodies with masses m_1 and m_2 at a distance *r*.

The constant γ is (up to a dimensional factor) denoted as **heavy mass** of a body. Equation (5.52) then implies the **equivalence of heavy and inertial mass.**

Kepler's 3rd law (5.48) with $c = \Gamma m_1 m_2$ then reads as:

$$T^2 = \frac{4\pi^2 m_1 m_2}{c(m_1+m_2)} a^3 = \frac{4\pi^2}{\Gamma(m_1+m_2)} a^3. \tag{5.54}$$

The ratio T^2/a^3 therefore is (practically) constant for all planets since $m_{planet} \ll m_{sun}$.

5.2.3 Equivalence Principle

Due to the **equivalence of inertial and heavy mass** (5.52) the force acting on a body of mass *m* in the earth's gravity field is

$$\overrightarrow{F} = m\overrightarrow{g}, \tag{5.55}$$

where the **gravitational field strength** $\overrightarrow{g}$ is independent of the properties of the body under consideration. Therefore, every body experiences the same acceleration at a certain place

$$\overrightarrow{a} = \overrightarrow{g}. \tag{5.56}$$

This result has an important **consequence**:

If an observer notices that different (electrically neutral) bodies at the same place experience the same acceleration $\overrightarrow{g}$, he can interpret this in two ways:

1. The system is an inertial system Σ and is located in a gravitational field that gives the same acceleration $\overrightarrow{g}$ for every body.
2. The observed bodies are free with respect to some inertial system Σ_k, but the observer system is located in an accelerated system Σ'. If its acceleration is $\overrightarrow{a}_0$, the acceleration $\overrightarrow{a}'_k$ measured relative to Σ' is connected with the acceleration $\overrightarrow{a}_k$ with respect to Σ_k by:

   $$\overrightarrow{a}'_k = \overrightarrow{a}_k - \overrightarrow{a}_0. \tag{5.57}$$

   If the bodies under consideration are free, $a_k = 0$, then they experience an acceleration relative to the observer in Σ' given by $\overrightarrow{a}'_k = -\overrightarrow{a}_0$. The experimental finding thus can also be explained with $\overrightarrow{a}_0 = -\overrightarrow{g}$.

Conclusion: An observer cannot determine whether his laboratory is in a homogeneous gravity field or in an accelerated reference system. This **equivalence principle** is the **basis of the theory of general relativity.**

5.2.4 Examples

(i)

Weightlessness in an earth satellite

(ii)

Minimum velocity for leaving the earth's gravitational field

According to Sect. 5.1.3 the **escape condition** (limiting case of the parabola!) is given by

$$E = \frac{1}{2}\mu v^2 - \Gamma \frac{mM_E}{R_E} = 0, \tag{5.58}$$

where R_E is the earth's radius, M_E the earth's mass and m the mass of the considered body; μ is the corresponding reduced mass, which may be replaced by m as long as $m \ll M_E$; v is the relative velocity of the body to the earth. From equation (5.58) we get for the **escape velocity**

$$v_F = \sqrt{\frac{2\Gamma M_E}{R_E}} \approx 10^4 \frac{\text{m}}{\text{sec}} \tag{5.59}$$

regardless of the mass of the body as long as $m \ll M_E$.

5.2.5 Gravitational Field of a Static Mass Distribution

A mass m' at position $\vec{r} = 0$ exerts the force to another mass located at $\vec{r} \neq 0$

$$\vec{F} = m\vec{g} \tag{5.60}$$

with

$$\overrightarrow{g(\vec{r})} = -\frac{\Gamma m'}{r^2}\frac{\vec{r}}{r} \; . \tag{5.61}$$

Interpretation: The mass m' creates a **gravitational field** at the position $\vec{r}$, whose strength (**gravity field strength**) is determined by $\overrightarrow{g(\vec{r})}$. The field strength $\vec{g}$ is a vector function that assigns a triple of real numbers to every point in space $\vec{r}$, i.e. $g_x(\vec{r}), g_y(\vec{r}), g_z(\vec{r})$, which during rotations behave like the components of a vector. Here $\overrightarrow{g(\vec{r})}$ shows always in the direction of the coordinate origin.

The potential energy corresponding to the force (5.60) is

$$U(\overrightarrow{r}) = m\phi(\overrightarrow{r}) \tag{5.62}$$

with

$$\phi(\overrightarrow{r}) = -\frac{\Gamma m'}{r} \ . \tag{5.63}$$

The quantity $\phi(\overrightarrow{r})$ is called the **potential** belonging to $\overrightarrow{g}$. Knowing $\phi(\overrightarrow{r})$ one can calculate $\overrightarrow{g(r)}$ via:

$$\overrightarrow{g} = -\operatorname{grad}\phi \ . \tag{5.64}$$

The function $\phi(\overrightarrow{r})$ describes a **scalar field** which assigns a real number to every point in space.

We can visualize the gravitational field of a resting mass point by its **field lines**: The tangent to a field line gives the direction of the force at each point $\overrightarrow{r}$ and the density of the field lines is a measure of the magnitude of the force. In the case of an individual mass point the associated field is always directed radially. The surfaces of constant potential then are spherical surfaces, whose common center lies at the origin of the coordinate system (see Fig. 5.8).

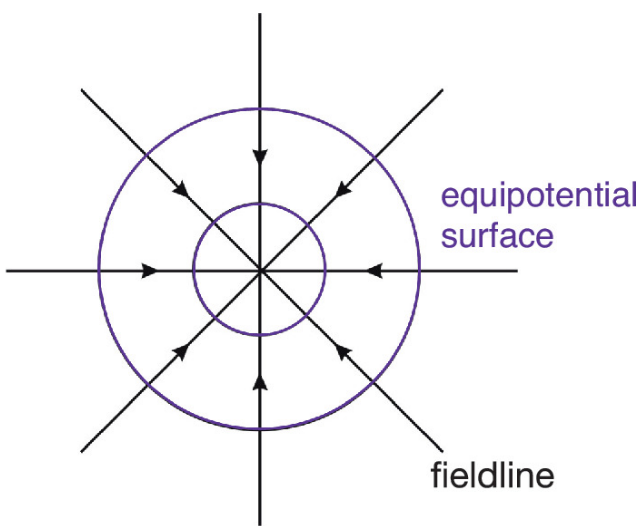

Fig. 5.8 Gravitational field lines and equipotential surfaces in case of a single mass located in the center

General statement:

When moving any test mass within an equipotential surface the potential ϕ does not change,

$$d\phi = \frac{\partial\phi}{\partial x}dx + \frac{\partial\phi}{\partial y}dy + \frac{\partial\phi}{\partial z}dz = (\mathrm{grad}\,\phi)\cdot\overrightarrow{\mathrm{dr}} = -\overrightarrow{\mathrm{g}}\cdot\overrightarrow{\mathrm{dr}} = 0\,. \qquad (5.65)$$

Since $\overrightarrow{dr} \neq 0$, it follows that $\overrightarrow{g}$ is perpendicular to the equipotential surfaces. This applies to every field whose field strength can be written as the gradient of a scalar field.

Of practical significance is the application to (discrete or continuous) mass distributions. According to the superposition principle (Chap. 1.2)

the gravitational field strength, generated by *N* mass points m_i at the positions $\overrightarrow{r_i}$, is given by:

$$\overrightarrow{g}(\overrightarrow{r}) = -\Gamma\sum_{i=1}^{N} m_i\frac{(\overrightarrow{r}-\overrightarrow{r_i})}{|\overrightarrow{r}-\overrightarrow{r_i}|^3}\ , \qquad (5.66)$$

or the potential ϕ by:

$$\phi(\overrightarrow{r}) = -\Gamma\sum_{i=1}^{N}\frac{m_i}{|\overrightarrow{r}-\overrightarrow{r_i}|}\ . \qquad (5.67)$$

For a continuous mass distribution the sums are replaced by integrals:

$$\overrightarrow{g}(\overrightarrow{r}) = -\Gamma\int\varrho(\overrightarrow{r'})\frac{(\overrightarrow{r}-\overrightarrow{r'})}{|\overrightarrow{r}-\overrightarrow{r'}|^3}\ d^3r' \qquad (5.68)$$

and

$$\phi(\overrightarrow{r}) = -\Gamma\int\frac{\varrho(\overrightarrow{r'})}{|\overrightarrow{r}-\overrightarrow{r'}|}\ d^3r'\ , \qquad (5.69)$$

where $\varrho(\overrightarrow{r'})$ denotes the mass density.

Example: Homogeneous sphere of radius R :

$$\varrho(\overrightarrow{r'}) = \begin{cases} \varrho_0 & r' \leq R \\ 0 & \text{else} \end{cases} \tag{5.70}$$

We carry out the volume integration within polar coordinates (see Fig. 5.9).

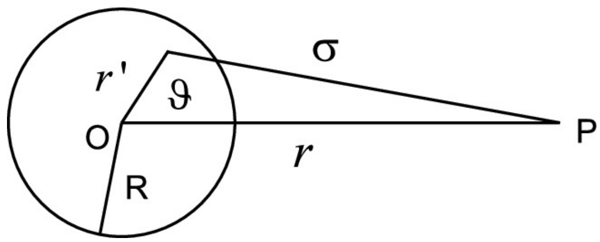

Fig. 5.9 Illustration of polar coordinates for the volume integral in case of a homogenous sphere

Using

$$\sigma^2 = (\vec{r} - \vec{r'})^2 = r'^2 + r^2 - 2rr' \cos\vartheta \quad ; \quad \frac{d\sigma^2}{d\vartheta} = 2\sigma\frac{d\sigma}{d\vartheta} = 2rr' \sin\vartheta;$$

$$2\sigma = 2rr' \sin\vartheta \tfrac{d\vartheta}{d\sigma}. \tag{5.71}$$

we find

$$\phi(\vec{r}) = -\Gamma\varrho_0 \textstyle\int_0^R \int_0^\pi \int_0^{2\pi} \frac{dr'r'd\vartheta\ r'\sin\vartheta d\varphi}{\sigma} \tag{5.72}$$

$$= -\Gamma\varrho_0 \int_0^R \int_0^\pi \int_0^{2\pi} dr'd\vartheta d\varphi\ \frac{r'^2 \sin\vartheta}{rr' \sin\vartheta} \frac{d\sigma}{d\vartheta} = -\frac{2\pi\Gamma\varrho_0}{r} \int_0^R \int_{\sigma_{\min}}^{\sigma_{\max}} r'dr'd\sigma.$$

Case 1: $r > R$ ($\sigma_{max} = r + r', \sigma_{min} = r - r'$)

$$\phi(\vec{r}) = -\tfrac{2\pi\Gamma\varrho_0}{r} (\textstyle\int_0^R r'(r + r' - (r - r'))\ dr' = -\frac{\Gamma}{r} \cdot \frac{4\pi}{3}\varrho_0 R^3 = -\frac{\Gamma M}{r}. \tag{5.73}$$

Only the total mass M and the distance r determine $\phi(r)$.

Case 2: $r < R$ For the integration we distinguish: i) $r > r'$, i.e. $\sigma_{max} = r + r', \sigma_{min} = r - r'$ and ii) $r' > r$, i.e. $\sigma_{max} = r + r', \sigma_{min} = r' - r$. Elementary integration gives:

$$\phi(\vec{r}) = -\frac{2\pi\Gamma\varrho_0}{r}\left(\int_0^r r'(r + r' - (r - r'))\ dr' + \int_r^R r'(r + r' - (r' - r))\ dr'\right)$$

$$= -\tfrac{2\pi\Gamma\varrho_0}{r}\left(\int_0^r\ 2r'^2 dr' + \int_r^R r' 2r\ dr'\right) = -4\pi\Gamma\varrho_0\left[\tfrac{R^2}{2} - \tfrac{r^2}{6}\right]. \tag{5.74}$$

The gravitational field $\vec{g}(\vec{r})$ then follows as negative gradient of $\phi(\vec{r})$, i.e. $\vec{g}(\vec{r}) = -\vec{\nabla}\phi(\vec{r})$.

5.3 Small Oscillations

5.3.1 The Linear Harmonic Oscillator

The equation of motion for a **linear harmonic oscillator** is:

$$m\ddot{x} = -kx\ , \qquad k > 0, \tag{5.75}$$

or with

$$\omega_0^2 = \tfrac{k}{m}, \tag{5.76}$$

$$\ddot{x} + \omega_0^2 x = 0. \tag{5.77}$$

The general (real) solution to the differential equation (5.77) is:

$$x = A_1\ \cos\ \omega_0 t + A_2\ \sin\ \omega_0 t \tag{5.78}$$

or

$$x = C\ \sin\ (\omega_0 t + \delta). \tag{5.79}$$

The general solution contains 2 integration constants A_1 and A_2 or C and δ. In (5.79) δ gives the phase of the oscillation at time $t = 0$; the amplitude C is linked to the energy which can be found as follows: the potential energy of the oscillator is

(5.80)

$$U(x) = \frac{1}{2}kx^2$$

such that the energy is given by

$$E = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2 = \text{const.} \tag{5.81}$$

or (with $k = m\omega_0^2$)

$$E = \frac{C^2}{2}\left\{m\omega_0^2 \cos^2\left(\omega_0 t + \delta\right) + k \sin^2\left(\omega_0 t + \delta\right)\right\} = \frac{kC^2}{2}. \tag{5.82}$$

At the inversion points $x = \pm C$ the kinetic energy $T = 0$ while the potential energy $U(\pm C)$ is maximum. In the equilibrium position ($x = 0$) the potential energy $U = 0$ and the kinetic energy is maximum (see Fig. 5.10).

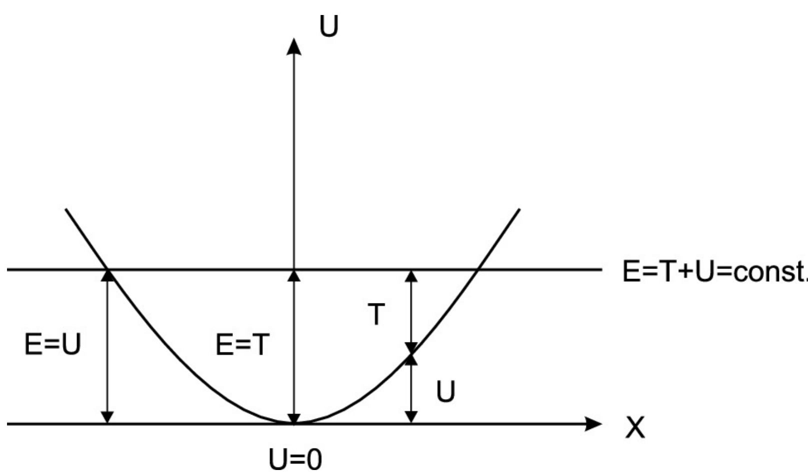

Fig. 5.10 Energy balance in case of a harmonic oscillator

Example: Thread pendulum (Fig. 5.11).

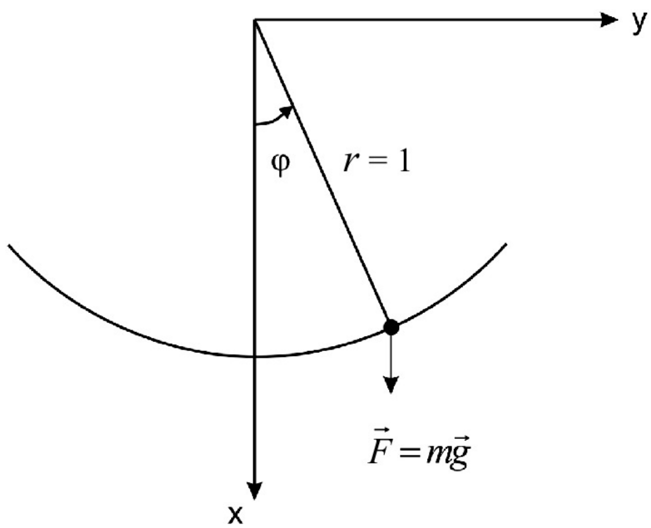

Fig. 5.11 Coordinates in case of a thread pendulum

The change in angular momentum is given by

$$\frac{d}{dt} l_z = \frac{d}{dt}(ml^2\dot{\varphi}) = (\overrightarrow{r} \times \overrightarrow{F})_z = -mgl \sin \varphi \tag{5.83}$$

or

$$\ddot{\varphi} + \frac{g}{l} \sin \varphi = 0. \tag{5.84}$$

For small deflections, $\sin \varphi \approx \varphi$, we get

$$\ddot{\varphi} + \omega_0^2 \varphi = 0 \text{ with } \omega_0^2 = \frac{g}{l}. \tag{5.85}$$

For larger pendulum deflections one obtains an **anharmonic vibration**.

5.3.2 Damped Oscillator

We extend the equation of motion (5.77) to:

$$\boxed{\ddot{x} + \omega_0^2 x + 2\beta\dot{x} = 0, \qquad \beta > 0,} \tag{5.86}$$

where the velocity-dependent term ($2\beta\dot{x}$) describes damping. With the solution $x(t) = e^{\lambda t}$ we get by insertion into (5.86):

$$\lambda^2 + \omega_0^2 + 2\beta\lambda = 0 \tag{5.87}$$

with the two solutions

$$\lambda_{1,2} = -\beta \pm \sqrt{\beta^2 - \omega_0^2}. \tag{5.88}$$

The general solution of (5.86) is a linear combination of the **basic solutions** $e^{\lambda_1 t}$ and $e^{\lambda_2 t}$. For the further discussion the following cases must be distinguished:

(i) $\beta < \omega_0$ **(weak damping)** With

$$\sqrt{\beta^2 - \omega_0^2} = i\omega \tag{5.89}$$

we can write the general solution as

$$x(t) = \left(A_1 e^{i\omega t} + A_2 e^{-i\omega t}\right) e^{-\beta t} \tag{5.90}$$

with the integration constants A_1 and A_2, or as a real function:

$$\boxed{x(t) = c e^{-\beta t} \sin\left(\omega t + \delta\right).} \tag{5.91}$$

This equation describes a damped oscillation (see Fig. 5.12).

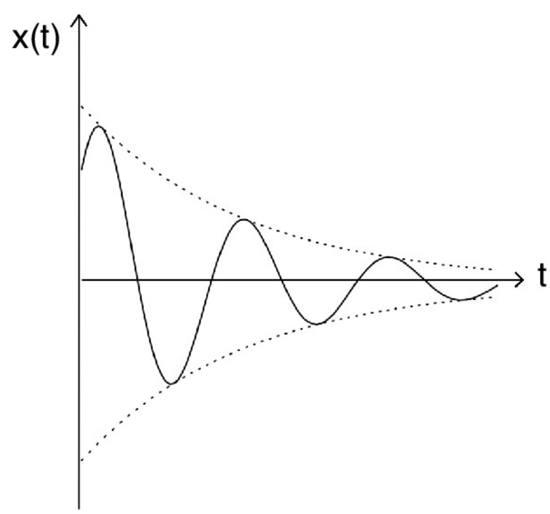

Fig. 5.12 Time dependence of the amplitude $x(t)$ in case of a weakly damped oscillator

(ii) $\beta = \omega_0$ **(critical damping)** In this case $\lambda_1 = \lambda_2$ and the approach $x = e^{\lambda t}$ only provides one of the two basic solutions. The second basic solution turns out to be

$$x(t) = te^{-\beta t}. \tag{5.92}$$

The general solution in the **aperiodic limit** takes the form:

$$\boxed{x(t) = A_1 e^{-\beta t} + A_2 t e^{-\beta t}.} \tag{5.93}$$

(iii) $\beta > \omega_0$ **(strong damping)** We set

$$\sqrt{\beta^2 - \omega_0^2} = \gamma > 0 \tag{5.94}$$

and obtain the general solution:

$$\boxed{x(t) = (A_1 e^{-\gamma t} + A_2 e^{\gamma t}) e^{-\beta t}.} \tag{5.95}$$

In this case we get an aperiodic motion; for $\beta > \gamma$ the amplitude $x(t) \to 0$ for large t.

Energy balance: Multiplying (5.86) by $m\dot{x}$ gives:

$$\boxed{\tfrac{d}{dt}\left(\tfrac{m}{2}\dot{x}^2 + \tfrac{k}{2}x^2\right) = -2\,m\beta\dot{x}^2 < 0.} \tag{5.96}$$

The oscillator is constantly losing energy due to friction ($\sim \beta$).

5.3.3 Forced Oscillations; Resonance

We now consider a damped harmonic oscillator driven by an external force $f(t)$ described by the equation of motion:

$$\boxed{\ddot{x} + \omega_0^2 x + 2\beta\dot{x} = \tfrac{1}{m} f(t).} \tag{5.97}$$

The general solution is composed of the general solution of the homogeneous equation and a specific solution of the inhomogeneous equation; the latter we determine for the important special case of a periodic force,

$$\tfrac{1}{m} f(t) = f_0 \cos \omega t. \tag{5.98}$$

Choosing

$$x(t) = \xi \cos(\omega t - \varphi) \tag{5.99}$$

we obtain from (5.97):

$$\xi\big(\left(\omega_0^2 - \omega^2\right) \cos(\omega t - \varphi) - 2\beta\omega \sin(\omega t - \varphi)\big) = f_0 \cos \omega t. \tag{5.100}$$

After squaring (5.100) and using the addition theorems

$$\cos(\alpha - \beta) = \cos\alpha \cos\beta + \sin\alpha \sin\beta$$

$$\sin(\alpha - \beta) = \sin\alpha \cos\beta - \cos\alpha \sin\beta \tag{5.101}$$

we get for the phase φ:

$$\boxed{\tan\varphi = \frac{2\beta\omega}{\omega_0^2 - \omega^2}} \tag{5.102}$$

and for the amplitude

$$\boxed{\xi = \frac{f_0}{\sqrt{\left(\omega^2 - \omega_0^2\right)^2 + 4\beta^2\omega^2}}.} \tag{5.103}$$

In addition to the special solution to the inhomogeneous equation, there is also the general solution of the homogeneous equation, i.e. a free damped oscillation. Due to the factor $e^{-\beta t}$ this part decays in time and for long times only the inhomogeneous solution remains as a **stationary solution** independent of the initial conditions.

The amplitude ξ and phase φ of the stationary solution have the following form as a function of ω (see Fig. 5.13):

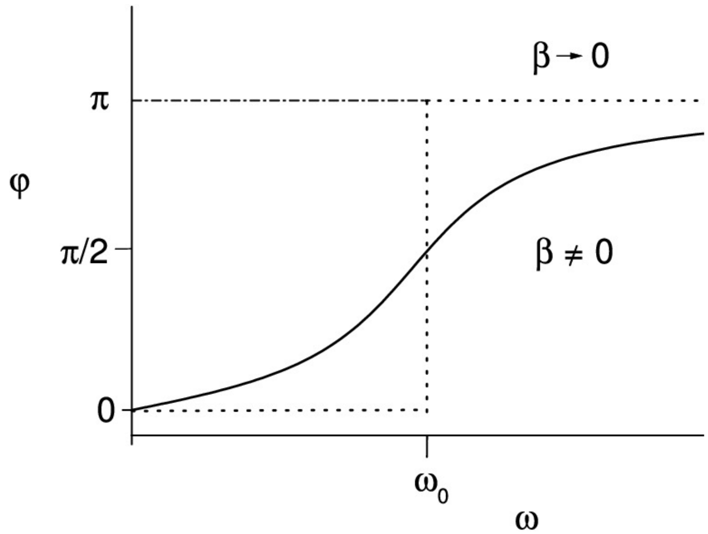

Fig. 5.13 The phase $\varphi(\omega)$ for the driven oscillator

For small frequencies ω the system follows the external force (practically) without delay: $\varphi \to 0$ for $\omega \to 0$. With increasing ω the phase φ increases, reaches $\pi/2$ for $\omega = \omega_0$, where the frequency of the external force is equal to the **natural frequency** ω_0 of the oscillator, and approaches the value π for $\omega \to \infty$, where the osillator is in antiphase to the external force.

For the special case $\beta \to 0$, φ changes suddenly from 0 to π for $\omega = \omega_0$ (dashed line in Fig. 5.13).

The amplitude ξ has the value f_0/ω_0^2 for $\omega = 0$. If $\omega_0^2 > 2\beta^2$, ξ grows with increasing frequency ω, reaches a maximum for $\omega_a = \sqrt{\omega_0^2 - 2\beta^2} \leq \omega_0$ and then approaches monotonically towards zero (see Fig. 5.14).

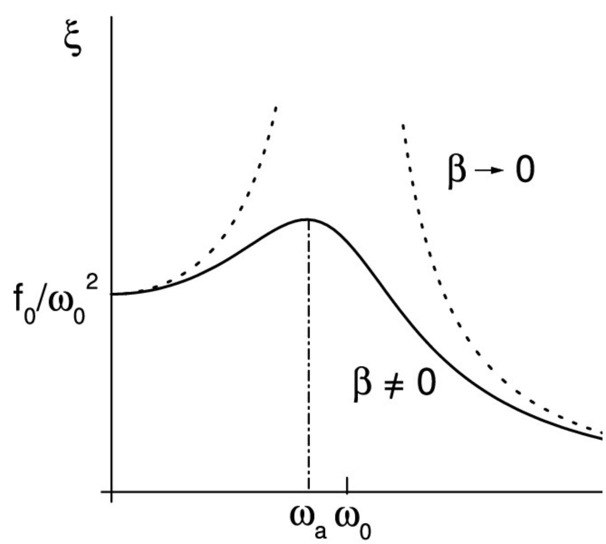

Fig. 5.14 The amplitude $\xi(\omega)$ for the driven oscillator

For strong damping, $2\beta^2 > \omega_0^2$, no maximum is formed; ξ tends towards zero as ω increases, starting at f_0/ω_0^2 for $\omega = 0$.

Of particular importance is the frequency $\omega = \omega_0$. There the phase φ passes the value $\pi/2$ and the work done by the external force becomes maximum (**energy resonance**).

Proof We calculate the average work done on the oscillator by the external force during time $T = 2\pi/\omega$:

$$W_f = \frac{1}{T}\int_0^T f(t)\ \dot{x}\ dt = \frac{mf_0}{T}\int_0^T\ \dot{x}(t)\ \ \cos\left(\omega t\right)\ dt, \tag{5.104}$$

where

$$\dot{x}(t) = -\xi\omega\sin\left(\omega t - \varphi\right). \tag{5.105}$$

Result

$$W_f = \frac{\beta m f_0^2\omega^2}{\left(\omega^2-\omega_0^2\right)^2+(2\beta\omega)^2}. \tag{5.106}$$

From

$$\frac{d}{d\omega}W_f(\omega) = \frac{d}{d\omega}\frac{\beta m f_0^2\omega^2}{\left(\omega^2-\omega_0^2\right)^2+(2\beta\omega)^2} = 0 \tag{5.107}$$

one finds that the average energy transferred to the oscillator W_f has a maximum for $\omega = \omega_0$.

The supplied energy W_f exactly compensates for the energy which the oscillator loses due to the damping—averaged over the period *T*- (5.96), i.e.

$$W_\beta = \frac{1}{T}\int_0^T \frac{dE}{dt}\ dt = -\frac{2\,m\beta}{T}\int_0^T \dot{x}^2\ dt = -W_f. \tag{5.108}$$

Examples:

1. **Ionic crystals**, e.g. NaCl

   If a light wave falls on such a crystal the oscillating electric field of the light wave generates a vibration of the positively charged ions relative to the negatively charged ions. The crystal absorbs energy from the light wave; the energy absorption of the crystal is maximum when the frequency

ω of the light coincides with the natural frequency ω_0 of the crystals vibration.

2.
By tuning an **electrical resonant circuit** one can adjust the natural frequency ω_0 of a radio to the frequency ω of the radio wave of a specific station. The receiver then dominantly absorbs radio waves of the desired station.

3.
Microwave oven
By tuning the frequency of the microwave ω_0 resonant vibrations of H_2O molecules are excited; the absorbed vibration energy is converted into thermal energy by interactions.

5.3.4 Coupled Harmonic Oscillations

Simple example: 2 coupled strings (Fig. 5.15).

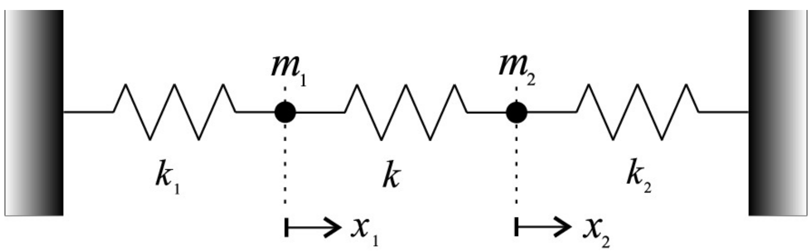

Fig. 5.15 Two particles of mass m_1 and m_2 are coupled by a string with strength k and attached to the outer walls by strings of strength k_1 and k_2

Two particles with masses m_1 and m_2, which can move only in a single dimension (x-axis), are coupled by an attractive force (k), which is proportional to the difference between the deflections from the rest position ($x_1 = 0$, $x_2 = 0$). The particles are also attached to their position by spring forces (k_1, k_2) to their resting positions. Then the equations of motion are:

$$m_1\ddot{x}_1 = -k_1x_1 - k(x_1 - x_2) \tag{5.109}$$

$$m_2\ddot{x}_2 = -k_2x_2 - k(x_2 - x_1). \tag{5.110}$$

The terms $-k_1x_1$ and $-k_2x_2$ are “external forces” but $k_{12} = -k(x_1 - x_2) = -k_{21}$ is an “internal force”, for which the actio $=$ reactio principle applies.

To solve the equations of motion we transform to:

$$\ddot{x}_1 + \omega_1^2 x_1 = \frac{k}{m_1} x_2 \tag{5.111}$$

$$\ddot{x}_2 + \omega_2^2 x_2 = \frac{k}{m_2} x_1 \tag{5.112}$$

with

$$\omega_i^2 = \frac{k+k_i}{m_i} \quad ; \quad i = 1.2 \tag{5.113}$$

Structure of the problem:

For $k = 0$ we have 2 decoupled oscillators; for $k \neq 0$ rhe right sides of (5.111) and (5.112) describe the coupling.

We continue to consider the simplified case:

$$m_1 = m_2 = m \; ; \; k_1 = k_2 = k_0 \to \omega_1 = \omega_2 = \omega_0 \tag{5.114}$$

leading to:

$$\ddot{x}_1 + \omega_0^2 x_1 = \frac{k}{m} x_2 \tag{5.115}$$

$$\ddot{x}_2 + \omega_0^2 x_2 = \frac{k}{m} x_1. \tag{5.116}$$

With the Ansatz

$$x_1 = a_1 \cos \omega t \quad ; \quad x_2 = a_2 \cos \omega t \tag{5.117}$$

we obtain

$$(\omega_0^2 - \omega^2) a_1 - \frac{k}{m} a_2 = 0 \tag{5.118}$$

and

$$-\frac{k}{m} a_1 + (\omega_0^2 - \omega^2) a_2 = 0. \tag{5.119}$$

In order to find non-trivial solutions for the unknown coefficients a_1 and a_2 of the linear system of equations, the determinant of the coefficients must vanish:

(5.120)

$$\begin{vmatrix} (\omega_0^2 - \omega^2) & -\frac{k}{m} \\ -\frac{k}{m} & (\omega_0^2 - \omega^2) \end{vmatrix} = 0$$

thus:

$$\left(\omega_0^2 - \omega^2\right)^2 = \tfrac{k^2}{m^2}. \tag{5.121}$$

The solutions are:

1. $\omega_a = \sqrt{k_0 + 2k/m}$; this leads to:

$$a_1 = -a_2, \tag{5.122}$$

   i.e. the particles oscillate in anti-phase (**antisymmetric vibration**).

2. $\omega_s = \sqrt{k_0/m}$: In this case we get a **symmetric oscillation**,

$$a_1 = a_2, \tag{5.123}$$

   i.e. the spring *k* is not vibrating at all and therefore the particles oscillate at the undisturbed frequency $\omega = \sqrt{k_0/m}$, as in case of no coupling.

In case (1.), however, the spring *k* during the vibration is stretched or pressed together. The general solution is a superposition of both solutions and reads:

$$x_1 = A_s \cos\left(\omega_s t + \alpha_s\right) + A_a \cos\left(\omega_a t + \alpha_a\right), \tag{5.124}$$

$$x_2 = A_s \cos\left(\omega_s t + \alpha_s\right) - A_a \cos\left(\omega_a t + \alpha_a\right). \tag{5.125}$$

It contains $2 \cdot 2 = 4$ free constants (A_s, A_a, α_s, α_a) corresponding to the number of degrees of freedom of the system.

The vibration types found above suggest to introduce **normal coordinates**:

$$\boxed{\begin{aligned} q_s &= x_1 + x_2 && (5.126) \\ q_a &= x_1 - x_2. && (5.127) \end{aligned}}$$

The variables q_s, q_a then follow decoupled equations of motion,

$$\ddot{q}_s + \left(\omega_0^2 - \tfrac{k}{m}\right) q_s = 0, \tag{5.128}$$

$$\ddot{q}_a + \left(\omega_0^2 + \tfrac{k}{m}\right) q_a = 0, \tag{5.129}$$

as can easily be found by insertion into (5.115) and (5.116). Accordingly one finds for the energy:

$$E = \frac{m}{2}\dot{x}_1^2 + \frac{k_0}{2}x_1^2 + \frac{m}{2}\dot{x}_2^2 + \frac{k_0}{2}x_2^2 + \frac{k}{2}(x_1 - x_2)^2$$

$$= \tfrac{m}{4}\dot{q}_s^2 + \tfrac{k_0}{4}q_s^2 + \tfrac{m}{4}\dot{q}_a^2 + \tfrac{k_0+2k}{4}q_a^2. \tag{5.130}$$

The method for decoupling vibrations—as outlined above—by introduction of normal coordinates is generally possible in the harmonic approximation.

Example: Vibrations of molecules and crystals.

In summarizing this chapter we have presented important applications of Newtonian mechanics for central forces, where the potential U only depends on the magnitude of the relative distance $|\overrightarrow{r_1} - \overrightarrow{r_2}|$ between two mass points. In this case the conservation of momentum, angular momentum and energy holds which drastically reduces the number of free degrees of freedom. An important case are $1/r^2$-forces, which holds for Coulomb and gravitational forces; we have classified the trajectories according to their energy and derived Kepler's laws for the motion of planets. In extension the law of gravity has been formulated and gravity fields been introduced for static mass distributions. In addition the dynamics of a linear oscillator was discussed and the solutions have been computed from the equations of motion also in case of additional frictional forces. The case of a damped oscillator, that is driven by an external periodic force, has lead to the formation of resonances that have been analysed with respect to the energy balance. In addition the problem of coupled harmonic oscillations was addressed that is solved by introducing 'normal' coordinates which decouple the equations of motion.

W. Cassing, *Theoretical Physics Compact I*
https://doi.org/10.1007/978-3-031-95512-9_6

6. Relativistic Mechanics

Wolfgang Cassing[1]
(1) University of Giessen, Gießen, Hessen, Germany

So far we have introduced classical Newton mechanics which, however, has different transformation properties than Maxwell's equations for electrodynamics. This incompatibility has been solved in **Einstein's special theory of relativity**: we have to replace the **Galilei transformation** between inertial systems by the **Lorentz transformation** that keeps the velocity of light *c* invariant in all inertial systems. We will derive the Lorentz transformation explicitly (in a simple case) and discuss its implications: Lorentz contraction, time dilation, simultaneity in moving systems as well as causality and the limiting velocity of signals. Some mathematical aspects of the Lorentz group of transformations will be discussed and Lorentz scalars, four-vectors and Lorentz tensors are identified as well as corresponding physical quantities like four-current densities. We close the discussion of relativistic dynamics by introducing the energy-momentum four-vector, which is conserved in all four components for closed systems and discuss scattering problems. As an example the important problem of Compton scattering of a photon on a resting charge *q* is computed explicitly. The derivation of the Lorentz transformation of the force will finalize this chapter.

6.1 Special Relativity

6.1.1 Lorentz Transformations

Galilei's principle of relativity (see Sect. 3.1.2) is:

The basic laws of mechanics exist in all inertial systems and have the same form.

If there are two inertial systems Σ and Σ' linked together by a Galilean transformation (see Sect. 3.1.2)

$$\vec{r'} = \vec{r} - \vec{v}_0 t; \qquad t' = t, \tag{6.1}$$

we obtain for the velocities:

$$\vec{v'} = \vec{v} - \vec{v}_0. \tag{6.2}$$

The relationships (6.1) and (6.2) have to be used, if two inertial observers—moving with constant velocity $\vec{v}_0$ relative to each other—want to compare measurements.

Newton's equations of motion (as the basic laws of classical mechanics) are indeed invariant for Galilean transformations, since according to (6.1) and (6.2) we have for the acceleration

$$\overrightarrow{a}' = \overrightarrow{a} \tag{6.3}$$

and the mass in Newtonian mechanics is a property independent of the state of motion for a mass point. The conservation laws for energy, momentum and angular momentum are also Galilean invariant statements (see Sects. 4.3 and 4.4).

Galileo's principle of relativity is well proven for 'small' particle velocities. Difficulties arise, however,

(i) for 'fast moving' particles and

(ii) in the connection with electrodynamics, especially optics.

Consider a light source moving—in relation to the observer—with the velocity $\overrightarrow{v_0}$, then according to (6.2) the velocity of a light source emitting signal $c \pm v_0$, will depend on whether the light source and observer approach each other or remove from each other. Maxwell equations (see electrodynamics)–especially the wave equations in vacuum–then could only apply to a single reference system. All attempts (such as the Michelson experiment), to prove the existence of such an **absolutely resting** system, have clearly failed.

The right conclusion from this obvious problem was drawn by Albert Einstein. His **special theory of relativity** is based on 2 postulates:

(1) **The laws of nature are the same in all inertial systems.**

(2) **The velocity of light in the vacuum is the same in all inertial systems.**

Since the postulates (1.) and (2.) are not compatible with (6.1), (6.2), we have to find a new transformation rule for the transition from an inertial system Σ to another inertial system Σ'.

6.1.2 Derivation of the Lorentz Transformation

We consider two inertial systems Σ, Σ' moving with constant velocity $v = v_0$ (for simplicity in $x-$ direction) relative to each other. A light signal is emitted from the origin *O* of Σ at time *t*=0, where *O* just coincides with the origin O' of Σ'. According to Einstein's principle of relativity, 2 observers—in Σ and Σ'—must describe the propagation of the light signal according to the same laws. For the observer in Σ the signal propagates as a spherical wave originating in *O*, whose front has the distance $r = ct$ from *O* at time *t*. The wavefront is therefore determined by:

$$r^2 = x^2 + y^2 + z^2 = c^2t^2. \tag{6.4}$$

For the observer in Σ' the center of the spherical wave is in O'; for him the following relation applies instead of (6.4):

$$r'^2 = x'^2 + y'^2 + z'^2 = c^2t'^2. \tag{6.5}$$

The observations (6.4) and (6.5) are not compatible with (6.1) since it follows from (6.5) (with (6.1)):

$$(x - vt)^2 + y^2 + z^2 = c^2t^2, \tag{6.6}$$

which for $v \neq 0$ does not agree with (6.4). We now try to modify (6.1), (6.2) such that (6.4) and (6.5) merge by the new transformation.

The transformation we are looking for must be linear such that the force-free motion of a particle in the system Σ to any other inertial system Σ' is force-free: the trajectory $\vec{r} = \vec{v}t$+ const. in Σ must be a linear one in $\vec{r}\,'$ and t' when transforming to Σ'. Due to the homogeneity of space and time we can always choose Σ and Σ' in such a way that for *t*=0 the points *O* and O' coincide; the transformation then is homogeneous. For the case selected above

$$\vec{v} = \begin{pmatrix} v \\ 0 \\ 0 \end{pmatrix} \tag{6.7}$$

we can always choose the axes in Σ' such that the x'-axis constantly coincides with the *x*-axis due to the isotropy of space. For a point on the *x* axis with $y = 0 = z$ in Σ then $y' = 0 = z'$ also holds in Σ'. The transformation

$$(x, y, z, ct) \rightarrow (x', y', z', ct') \tag{6.8}$$

then separates such that

$$(x, ct) \rightarrow (x', ct') \tag{6.9}$$

and

$$(y, z) \rightarrow (y', z'). \tag{6.10}$$

By rotation around the *x* axis one then can always achieve that

$$y = \lambda y', \qquad z = \lambda z' \ . \tag{6.11}$$

Due to the equivalence of the systems Σ and Σ' we must have $\lambda = 1$, wich gives

$$y' = y; \qquad z' = z. \tag{6.12}$$

For the transformation (6.9) we assume a linear transformation in *x* and *t*:

$$x' = a_1x + a_2t; \qquad t' = a_3x + a_4t. \tag{6.13}$$

Since the origin O' of Σ' relative to Σ has the velocity *v*, it follows from

$$0 = a_1x + a_2t \tag{6.14}$$

immediately

$$a_2 = -a_1v. \tag{6.15}$$

Thus (6.13) turns to:

$$x' = a_1(x - vt); \qquad t' = a_3x + a_4t. \tag{6.16}$$

We determine the remaining coefficients a_1, a_3, a_4 from the requirement that (6.5) with (6.12), (6.16) should give (6.4). For

$$(a_1^2 - a_3^2 c^2)x^2 + y^2 + z^2 = 2(a_1^2 v + c^2 a_3 a_4)xt + (c^2 a_4^2 - a_1^2 v^2)t^2 \tag{6.17}$$

to match (6.4) for all x, y, z, t the following conditions must hold:

$$a_1^2 - c^2 a_3^2 = 1; \qquad a_4^2 - \beta^2 a_1^2 = 1; \qquad a_1^2 v + c^2 a_3 a_4 = 0 \tag{6.18}$$

with the abbreviation

$$\beta = \tfrac{v}{c}. \tag{6.19}$$

The combination of the first two equations in (6.18) yields

$$c^2 a_3^2 a_4^2 = (a_1^2 - 1)(1 + \beta^2 a_1^2); \tag{6.20}$$

the 3rd equation in (6.18) (solved for $a_3 a_4$) gives:

$$a_1^4 \beta^2 = (a_1^2 - 1)(1 + \beta^2 a_1^2) = a_1^2 + a_1^4 \beta^2 - a_1^2 \beta^2 - 1, \tag{6.21}$$

thus

$$-a_1^2 + 1 + \beta^2 a_1^2 = 0 \qquad\qquad \to a_1^2 = \tfrac{1}{1-\beta^2}. \tag{6.22}$$

With (6.22) we obtain for (6.18):

$$a_4^2 = 1 + \tfrac{\beta^2}{1-\beta^2} = a_1^2; \qquad\qquad a_3^2 = \tfrac{a_1^2 - 1}{c^2} = \tfrac{\beta^2}{c^2(1-\beta^2)}. \tag{6.23}$$

The choice of the sign is still pending: For $\beta \to 0$ (6.12) and (6.13) should go over to (6.1), i.e.:

$$a_1 = a_4 = \tfrac{1}{\sqrt{1-\beta^2}}; \qquad\qquad a_3 = -\tfrac{\beta}{c\sqrt{1-\beta^2}}. \tag{6.24}$$

The Lorentz transformation then reads:

$$x' = \tfrac{x - vt}{\sqrt{1-\beta^2}}; \qquad y' = y; \qquad z' = z; \qquad t' = \tfrac{t - vx/c^2}{\sqrt{1-\beta^2}}. \tag{6.25}$$

The inversion

$$x = \tfrac{x' + vt'}{\sqrt{1-\beta^2}}; \qquad y = y'; \qquad z = z'; \qquad t = \tfrac{t' + vx'/c^2}{\sqrt{1-\beta^2}} \tag{6.26}$$

is obtained by replacing v by $-v$, i.e. by exchanging the systems Σ and Σ' (of equal rights).

6.1.3 Space-Time Diagrams

The connections between inertial systems can be summarized and displayed in space-time diagrams. Except for the coordinate $x_0 = ct$ let's consider another representative coordinate x_1. Points (x_0, x_1), or generally (x_0, x_1, x_2, x_3), in this diagram are called **events** or **world points**. The connection of two world points by a **world line** can be the path of a mass point or a light signal.

It is crucial for the representation of events in different inertial systems that the **world distance** of an event from the origin

$$\boxed{s^2 = c^2t^2 - r^2} \tag{6.27}$$

is invariant for Lorentz transformations (see (6.4), (6.5)). In the 2-dimensional representation in the x_0, x_1- plane the distance squared is

$$r^2 = x_1^2; \tag{6.28}$$

in general:

$$r^2 = x_1^2 + x_2^2 + x_3^2. \tag{6.29}$$

According to (6.4) the propagation of light, i.e. the world lines of photons, is characterized by

$$\boxed{s^2 = 0.} \tag{6.30}$$

In the 2-dimensional representation (6.30) reduces to the two straight lines

$$x_1 = \pm x_0; \tag{6.31}$$

if we add another position coordinate x_2, we get (from (6.31)) a cone by rotating around the x_0 axis (**light cone**), in the general case we get a hypercone in 4 dimensions. Equation (6.30) describes for $x_0 < 0$ a light signal arriving at the origin (0, 0) and for $x_0 > 0$ a light signal emitted by (0, 0).

The light cone divides the **Minkowski space** in 2 areas (Fig. 6.1) for

$$s^2 > 0 \text{ and } s^2 < 0. \tag{6.32}$$

The area $s^2 > 0$ includes the **past**, $x_0 < 0$, from which an observer in (0, 0) can receive signals, and the **future**, $x_0 > 0$, into which it can send signals. World lines that we can physically realize always run in the area $s^2 > 0$, since *c* is the limiting velocity for the transport of matter or energy (see below).

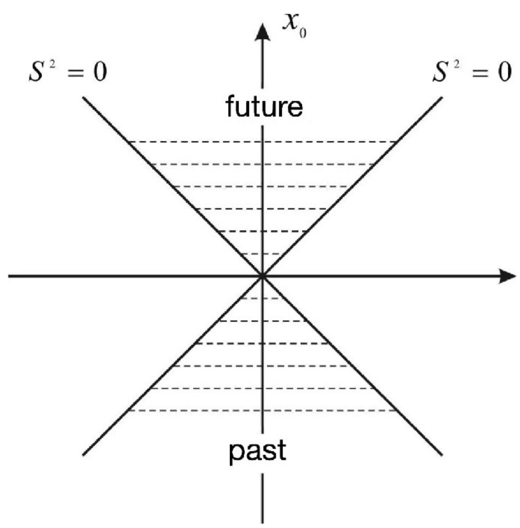

Fig. 6.1 Space-time diagram dividing past and future as well as time-like (dashed) and space-like areas

The area $s^2 < 0$ cannot be reached; we can neither get there (or send signals) nor receive signals from there. In this (**space-like**) region there could be particles, for which the velocity of light *c* forms a lower limit (**tachyons**). However, such speculations are without implications for the further formulation of relativistic mechanics.

Note: The division of past, future and space-like world points ($s^2 < 0$) is the same in every inertial frame since the separating light cone is Lorentz invariant!

In order to display 2 inertial systems Σ, Σ' in a Minkowski diagram we write (6.25) in the form

$$x'_0 = \gamma(x_0 - \beta x_1); \qquad x'_1 = \gamma(x_1 - \beta x_0) \tag{6.33}$$

using the abbreviation

$$\gamma = \frac{1}{\sqrt{1-\beta^2}}. \tag{6.34}$$

The x'_1 axis then consists of all points with x'_0 = 0; conversely, the x'_0 axis is determined by x'_1 = 0. It follows from (6.33):

$$0 = \gamma(x_0 - \beta x_1) \to x_0 = \beta x_1; \; 0 = \gamma(x_1 - \beta x_0) \to x_1 = \beta x_0. \tag{6.35}$$

Thus the axes in Σ' are straight lines through the origin (see Fig. 6.2); they are symmetrically to the light cone and are inclined by the angle α with respect to the axes in Σ with

$$\tan\alpha = \beta. \tag{6.36}$$

To define (time and length) units we use that the size s^2 is Lorentz-invariant. The intersection of the hyperbola (or the single-shell hyperboloid)

$$s^2 = -1 \tag{6.37}$$

with the (positive) x_1 or x_1' axis is the point (0, 1) in Σ or Σ' and defines the unit of length. The intersection of the hyperbola (or the double-shell hyperboloid)

$$s^2 = +1 \tag{6.38}$$

with the (positive) x_0 or x_0' axis is the point (1, 0) in Σ or Σ' which defines the time unit.

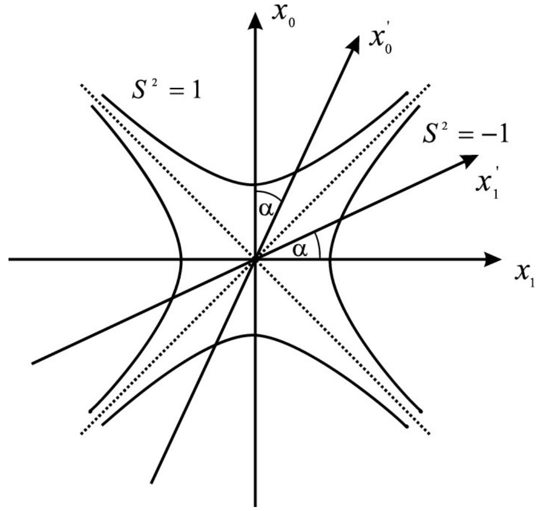

Fig. 6.2 Illustration of a Lorentz transformation in *x*-direction with velocity β. The x_0 and x_1 axes are tilted by the angle α defined by $\tan{(\alpha)} = \beta$ in Σ'

6.2 Consequences of the Lorentz Transformations

6.2.1 Addition of Velocities

A mass point moves with velocity $\overrightarrow{v}$ in Σ,

$$\overrightarrow{v} = \frac{d\overrightarrow{r}}{dt}. \tag{6.39}$$

We now look for the connection between $\overrightarrow{v}$ and the velocity of the mass point $\overrightarrow{v}'$ that another inertial observer found in Σ',

$$\overrightarrow{v}' = \frac{d\overrightarrow{r}'}{dt'}. \tag{6.40}$$

To this aim we form the differentials to (6.25):

$$dx' = \gamma(dx - vdt) = \gamma(v_x - v)dt; \quad dt' = \gamma(dt - \tfrac{v}{c^2}dx) = \gamma(1 - \tfrac{vv_x}{c^2})dt. \tag{6.41}$$

Then we obtain

$$v_y' = \frac{dy'}{dt'} = \frac{dy}{dt}\frac{\sqrt{1-\beta^2}}{(1-vv_x/c^2)} = v_y\frac{\sqrt{1-\beta^2}}{(1-vv_x/c^2)} \tag{6.42}$$

and as well

$$v'_z = \frac{dz'}{dt'} = \frac{dz}{dt}\frac{\sqrt{1-\beta^2}}{(1-vv_x/c^2)} = v_z\frac{\sqrt{1-\beta^2}}{(1-vv_x/c^2)}, \tag{6.43}$$

since $dy' = dy, dz' = dz$. On the other hand, for v'_x we get (6.41):

$$v'_x = \frac{dx'}{dt'} = \frac{dx'}{dt}\frac{\sqrt{1-\beta^2}}{(1-vv_x/c^2)} = \frac{v_x - v}{(1-vv_x/c^2)}. \tag{6.44}$$

Special case: For $v_y = v_z = 0$ the component v'_x becomes

$$v'_x = \frac{v_x - v}{(1-vv_x/c^2)}; \qquad v'_y = 0, \qquad v'_z = 0, \tag{6.45}$$

and we obtain for the limiting case

(i) $v \ll c$

$$v'_x = v_x - v \tag{6.46}$$

directly (6.2) and in the limit

(ii) $v \to c$

$$|v'_x| \to c, \tag{6.47}$$

thus showing that *c* plays the role of a limiting velocity.

6.2.2 Lorentz Contraction

We consider a rod of length l_0 in the system Σ which is at rest and (for simplicity) located in *x*-direction. The coordinates of the end points of the rod then are x_1, x_2 independent of time *t* in Σ and

$$l_0 = x_2 - x_1 \tag{6.48}$$

is the **rest length** of the rod. In order to calculate the length of the rod in a system Σ'—moving relative to Σ with the velocity *v* in *x*-direction—one has to consider the coordinates of the end points x'_1, x'_2 **simultaneously** in Σ', i.e. at a time $t'_1 = t'_2 = t'$; then the length

$$l' = x'_2 - x'_1 \tag{6.49}$$

is linked to l_0 according to (6.25) by:

$$l_0 = x_2 - x_1 = \gamma(x'_2 - x'_1) = \gamma l' \tag{6.50}$$

or

$$l' = \frac{l_0}{\gamma} < l_0, \tag{6.51}$$

since $\gamma > 1$. The observer moving relative to the rod in Σ' finds its length to be shorter than the rest length in Σ (**Lorentz contraction**). Perpendicular to the direction of motion the length measurements in Σ and Σ' give the same result.

On the other hand, if an observer in Σ measures a rod resting in Σ', he also finds a contraction according to the principle of relativity and not an elongation! The Lorentz contraction does not change the object **rod**, but only the different points of view of the observers in Σ and Σ'.

6.2.3 Simultaneity

We consider two events that occur in the inertial frame Σ in the points x_1 and x_2 with $x_1 \neq x_2$ at the same time $t_1 = t_2 = t$. After the Lorentz transformation (6.25) the two events in another inertial system Σ' then are not only spatially separated, $x_1' \neq x_2'$, but also in time $t_1' \neq t_2'$: The event occurring at time t and position x_1 in Σ is taking place in Σ' at the time

$$t_1' = \gamma(t - vx_1/c^2); \tag{6.52}$$

accordingly, the event occurring in Σ at the position x_2 and time t, at time

$$t_2' = \gamma(t - vx_2/c^2) \tag{6.53}$$

in Σ'. Thus

$$\boxed{\Delta t' = t_2' - t_1' \neq 0} \tag{6.54}$$

if $x_1 \neq x_2$. The simultaneous events in Σ are no longer simultaneous in Σ'.

Simultaneity can only be defined in a specific system and is lost when switching to another system! This implies that Newton's concept of an **absolute time** has to be abandoned.

6.2.4 Time Dilation

We consider a transmitter at position x in the system Σ, which sends out signals at a time difference

$$\Delta t = t_2 - t_1. \tag{6.55}$$

For an observer in a system Σ', which is moving with constant velocity v along the x-axis of Σ follows for the time interval between the signals (6.25)

$$\boxed{\Delta t' = t_2' - t_1' = \gamma \Delta t > \Delta t.} \tag{6.56}$$

The time $\Delta t'$ measured in Σ' is therefore longer than the **proper time** $\tau = \Delta t$ of the transmitter measured in Σ (**time dilation**). Observers in various inertial systems measure different time intervals but via (6.56) all calculate the same proper time τ. In analogy to the Lorentz contraction the time dilation is not a change in the object **transmitter**.

6.2.5 Causality and Limiting Velocity of Signals

The principle of causality states:

If an event A is the cause of another event B , then there cannot be an inertial system in which B occurs before A .

Otherwise by changing the reference system the temporal order of cause and effect would be reversed.

As a consequence of the causality principle, the velocity of light c in vacuum is an upper limit for the transmission of information in the form of energy transport (light signal) or mass transport (exchange of particles).

Explanation: A neutron may be created in the system Σ at point A (e.g. by the decay of an excited nucleus) and move from position A to position B, where it decays. Then, according to the principle of causality, there is no other inertial system Σ', where for an observer the neutron in B' decays before it is formed in A'.

We now assume that the neutron moves with velocity $v = \eta c$ with $\eta > 1$ and show that this contradicts the principle of causality: In Σ' one finds for the time interval $\Delta t'$ between formation and decay of the neutrons

$$\Delta t' = \gamma(\Delta t - v\Delta x/c^2), \tag{6.57}$$

if Σ' moves relative to Σ with velocity v along the $x-$ direction. Δt is the running time of the neutrons in Σ, Δx the corresponding distance,

$$\Delta x = \eta c \Delta t. \tag{6.58}$$

This will give:

$$\Delta t' = \gamma \Delta t(1 - \tfrac{\eta v}{c}), \tag{6.59}$$

and since we assumed $\eta > 1$, we can choose $v < c$ such that

$$(1 - \tfrac{\eta v}{c}) < 0. \tag{6.60}$$

In this case there would be a system Σ' in which $\Delta t' < 0$ but $\Delta t > 0$, i.e. in which the neutron in B' decays before it was created in A'!

Note: The considerations above do not exclude that 'geometrical' velocities $> c$ occur. For example, a light spot, emitted from a laser beam from the earth to the moon, may move with a velocity $> c$ over the lunar surface. This does not contradict the principle of causality because the path of the light spot on the moon is just the ensemble of impact points of individual light pulses, each of which travels the distance from earth to the moon with the velocity c. The velocity of the light spot is not the same as the transport of mass or energy on the lunar surface! Velocities $> c$ can also be achieved in the propagation of electromagnetic waves in dispersive media in form of **phase velocities** (see electrodynamics).

6.2.6 Examples and Explanations

Lifetime of muons

An example for time dilation is provided by the observation of muons ($\mu^{\pm}$), which are produced by cosmic radiation in the earth's atmosphere and are observed at the earth's surface. The muons are created between h_{min} = 10 km and h_{max} = 20 km above the earth's surface; their minimum running time is then

$$\Delta t = \tfrac{h_{min}}{c} \approx 30 \cdot 10^{-6} sec. = 30 \mu s. \tag{6.61}$$

However, the lifetime of a muon at rest is only $\tau \approx 2\ \mu s$, which corresponds to a maximum running distance of $c\tau \approx 600$ m! Consequently, muons created in the earth's atmosphere cannot reach the earth's surface at all according to Newtonian mechanics!

The apparent contradiction is resolved within the framework of Einstein's theory of relativity: The decay of muons is a structure property and therefore the lifetime τ comparable to the proper time of a clock. The lifetime in the rest frame τ therefore has to be distinguished from the time Δt measured by an observer on earth; equation (6.56) shows that for $\beta \approx 0.98$ the above values for τ and Δt are compatible with each other. On the other hand, the problem is solved from the perspective of the muon's rest system by the Lorentz contraction of the distance from the upper atmosphere to the earth's surface.

Lorentz contraction in the Minkowski diagram

We consider a unit scale at rest in Σ, which at time $t = 0$ may have the endpoints *O* and *A*. In the Minkowski diagram the scale moves perpendicular to the x_1 axis in positive x_0 direction. For an observer in Σ' the length of the scale is given by the distance OA', which is obviously shorter than the length unit OB' in Σ'. For an observer in Σ the latter appears shortened to the distance *OB* (see Fig. 6.3).

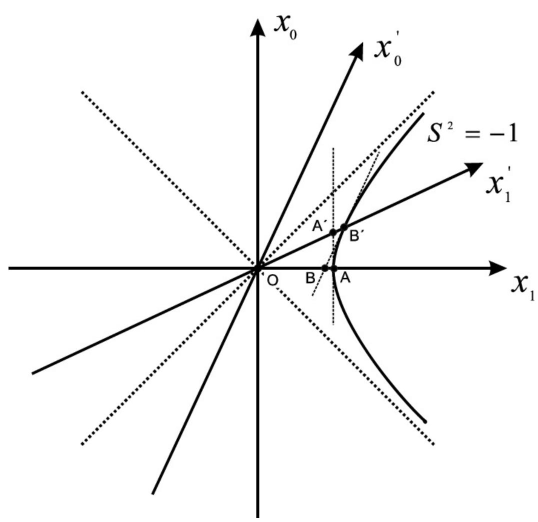

Fig. 6.3 For an observer in Σ' the length of the scale is given by the distance OA' while for an observer in Σ the latter appears shortened to the distance *OB*

6.3 Mathematical Aspects of Lorentz Transformations

In this section we will show that the basic equations of relativistic mechanics have the same form in all inertial frames (**covariance**) and thus obey the principle of relativity. Before, however, we will examine the mathematical structure of the Lorentz transformations.

6.3.1 Lorentz Group

First of all it will be shown that the Lorentz transformation is a complex orthogonal transformation in a 4-dimensional pseudo-euclidean vector space (**Minkowski space**). To this aim we introduce the following coordinates:

$$x_0 = ict, \qquad x_1 = x, \qquad x_2 = y, \qquad x_3 = z. \tag{6.62}$$

The square of the length of a space-time vector in different reference systems Σ and Σ' then can be written as:

$$\sum_{\mu=0}^{3} x_\mu^2 = \sum_{\mu=0}^{3} x_\mu'^2. \tag{6.63}$$

Comment:

The imaginary component x_0 might appear disturbing at first sight but this can be counterbalanced by a redefinition of the metric in the scalar product

$$\sum_{\mu=0}^{3} x_\mu x_\mu \to \sum_{\mu=0}^{3} x_\mu' \sum_{\nu=0}^{3} g_{\mu\nu} x_\nu' \tag{6.64}$$

with the new real component $x_0' = ct$ ($x_k' = x_k$ for k=1,2,3) and the pseudo-metric tensor

$$g_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{6.65}$$

with determinant $\det\left(g_{\mu\nu}\right) = -1$. An alternative to (6.65) is

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{6.66}$$

Both choices have been used in the literature and one has to take care about signs in the formulae presented according to the choice of the metric.

A **general Lorentz transformation**

$$x_\mu' = \sum_\nu a_{\mu\nu} x_\nu; \mu, \nu = 0, 1, 2, 3 \tag{6.67}$$

must keep the **length** of the vector (x_0, x_1, x_2, x_3) invariant:

$$\sum_{\mu=0}^{3} x_\mu^2 = \vec{r}^2 - c^2 t^2 = -s^2 = \text{const.} \tag{6.68}$$

In analogy to the 3-dimensional euclidean space this condition can be written as an orthogonality relation for the transformation coefficients $a_{\mu\nu}$:

$$\sum_\nu a_{\mu\nu}^T a_{\nu\lambda} = \delta_{\mu\lambda}, \tag{6.69}$$

where a^T is the transposed matrix to the matrix a. Equation (6.69) follows from:

$$\sum_\mu x_\mu'^2 = \sum_\mu \sum_{\nu\nu'} a_{\mu\nu} a_{\mu\nu'} x_\nu x_{\nu'} = \sum_{\nu\nu'} \{\sum_\mu a_{\nu\mu}^T a_{\mu\nu'}\} x_\nu x_{\nu'} = \sum_{\nu\nu'} \delta_{\nu\nu'} x_\nu x_{\nu'} = \sum_\nu x_\nu^2. \tag{6.70}$$

For a Lorentz transformation in x_1- direction with velocity $\beta = v/c$ the transformation matrix $a_{\mu\nu}$ has the special form

(6.71)

$$a_{\mu\nu} = \begin{pmatrix} \gamma & -i\gamma\beta & 0 & 0 \\ i\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

with $\gamma^2 = 1/(1-\beta^2)$.

The specification of the x_1-axis contained in (6.71) can be corrected by an additional orthogonal transformation in $\mathbb{R}^3$ in the form of a rotation. This possibility is based on the fact that Lorentz transformations form a **group**:

(1.) If we carry out 2 Lorentz transformations one after the other,

$$x'_\mu = \sum_\nu a_{\mu\nu} x_\nu; \qquad x''_\rho = \sum_\nu a'_{\rho\nu} x'_\nu; \qquad (\Sigma \to \Sigma' \to \Sigma''), \tag{6.72}$$

the result

$$x''_\rho = \sum_{\nu,\mu} a'_{\rho\nu} a_{\nu\mu} x_\mu = \sum_\mu a''_{\rho\mu} x_\mu; \qquad (\Sigma \to \Sigma''), \tag{6.73}$$

is again a Lorentz transformation because for the matrices a'', a' and a we have:

$$(a'')^T a'' = (a'a)^T (a'a) = a^T (a'^T a') a = a^T a = 1_4, \tag{6.74}$$

since by definition

$$a^T a = 1_4; \qquad (a')^T a' = 1_4 \tag{6.75}$$

with 1_4 denoting the 4×4 identity matrix. The connection between the elements of the group is therefore the (4×4) matrix multiplication.

(2.) The neutral element is the 1_4 matrix for Lorentz transformations with velocity *v* = 0.

(3.) For every transformation *a* there exists the inverse, since from (6.69) we have:

$$\det(a^T a) = (\det(a))^2 = 1, \tag{6.76}$$

thus

$$\det(a) \neq 0. \tag{6.77}$$

(4.) Since the matrix multiplication is associative, the associative law applies also to Lorentz transformations.

The orthogonal transformations in $\mathbb{R}^3$ (rotations and reflections) form a subgroup of the Lorentz group represented by

$$d_{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & d_{ik} \end{pmatrix} \tag{6.78}$$

with *i*, *k* = 1,2,3 and

(6.79)

$$\sum_{m=1}^{3} d_{im}^T d_{mj} = \delta_{ij}.$$

The general Lorentz transformation (6.67) with the condition (6.69) is obtained by combining (6.71) with (6.78), (6.79) and adding the **time inversion**

$$x'_i = x_i; \qquad x'_0 = -x_0; \qquad i = 1, 2, 3. \tag{6.80}$$

The Lorentz transformations therefore include: rotations in $\mathbb{R}^3$, space reflections and time inversion as well as the transition between inertial systems that move with a constant relative velocity towards each other.

When there is a translation in space or time, this does not change condition (6.63) because it only affects spatial and temporal distances.

To the group of **homogenous Lorentz transformations** (discussed above) we can therefore add translations in space and time and then get the 10-parameter **Poincaré group**, which has 3 parameters for spatial rotations, 3 parameters for Lorentz boosts with the velocity $\vec{v}$ and 4 parameters for space-time translations. **It is considered as the basic invariance group of physical systems**.

6.3.2 Lorentz Scalars, Vectors, Tensors

In analogy to the group of rotations we now define tensors with respect to the Lorentz group:

(1.)

Lorentz scalar

We denote a quantity Ψ a **Lorentz scalar**, if Ψ does not change for Lorentz transformations,

$$\Psi \to \Psi' = \Psi. \tag{6.81}$$

Examples are the electric charge, the mass squared M^2 (see Sect. 6.4) or the space-time distance squared s^2.

(2.)

Lorentz vector

We define a quantity A_μ to be a **Lorentz or four-vector**, if in Lorentz transformations its components A_μ $(\mu = 0, 1, 2, 3)$ transform the same as the components x_μ

$$A_\mu \to A'_\mu = \sum_{\nu=0}^{3} a_{\mu\nu} A_\nu. \tag{6.82}$$

Examples:

(i) The partial derivatives of a Lorentz scalar Ψ with respect to x_μ form the components of a four-vector because:

$$\frac{\partial \Psi'}{\partial x'_\mu} = \sum_\nu \frac{\partial \Psi}{\partial x_\nu} \frac{\partial x_\nu}{\partial x'_\mu} = \sum_\nu a_{\mu\nu} \frac{\partial \Psi}{\partial x_\nu} \tag{6.83}$$

using the inverse formula to (6.67):

$$x_\nu = \sum_\rho a_{\rho\nu} x'_\rho. \tag{6.84}$$

(ii)

The **4-divergence** of a four-vector is a Lorentz scalar:

$$\sum_\nu \frac{\partial A'_\nu}{\partial x'_\nu} = \sum_\nu \sum_{\mu,\mu'} a_{\nu\mu} a_{\nu\mu'} \frac{\partial A_\mu}{\partial x_{\mu'}} = \sum_\mu \frac{\partial A_\mu}{\partial x_\mu} \tag{6.85}$$

considering (6.69).

(iii)

Choosing the components of the four-vector according to (6.82)

$$A_\mu = \frac{\partial \Psi}{\partial x_\mu}, \tag{6.86}$$

we find (6.85):

$$\sum_\nu \frac{\partial^2}{\partial x_\nu^2} \Psi = \sum_\nu \frac{\partial^2}{\partial x_\nu'^2} \Psi', \tag{6.87}$$

i.e. the operator $\left(\triangle - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) = \sum_\mu \partial^2/\partial x_\mu^2$ thus is invariant for Lorentz transformations. Then for a four-vector with components A_μ the (wave) equation

$$\sum_\nu \frac{\partial^2}{\partial x_\nu^2} A_\mu = \left(\triangle - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right) A_\mu \tag{6.88}$$

transforms like the μ-th component of a four-vector (see electrodynamics).

(iv)

The scalar product of two four-vectors is a Lorentz scalar:

$$\sum_\mu A'_\mu B'_\mu = \sum_\mu \sum_{\nu,\rho} a_{\mu\rho} a_{\mu\nu} A_\rho B_\nu = \sum_\nu A_\nu B_\nu. \tag{6.89}$$

(3.)

Lorentz tensors of 2nd rank

Except for the scalars ($\equiv$ tensors of 0th rank) and the vectors ($\equiv$ tensors of 1st rank) we will encounter tensors of 2nd rank. They are defined as 4×4 matrices; their components $F_{\mu\nu}$ have the transformation property

$$\boxed{F'_{\mu\nu} = \sum_{\lambda,\sigma} a_{\mu\lambda} a_{\nu\sigma} F_{\lambda\sigma}.} \tag{6.90}$$

6.3.3 Four-Current Density

As an example for a four-vector we examine the transformation properties of the **sources** $\vec{j}$ and ρ of the electromagnetic field. The conservation of charge serves as a starting point:

$$\boxed{\nabla \cdot \vec{j} + \frac{\partial \rho}{\partial t} = 0.} \tag{6.91}$$

With the notations

$$j_0 = ic\rho; \qquad j_1 = j_x; \qquad j_2 = j_y; \qquad j_3 = j_z \tag{6.92}$$

we can write the continuity equation in four-notation as

$$\sum_\mu \frac{\partial}{\partial x_\mu} j_\mu = 0. \tag{6.93}$$

Because of charge invariance (6.93) holds in every inertial system since (6.93) is invariant for Lorentz transformations. Then—according to (6.85)—the components j_μ are the components of a four-vector (**four-current density**).

6.4 Relativistic Dynamics

Newton's equations of motion are invariant with respect to Galilei transformations but not with respect to Lorentz transformations (cf. Sect. 6.1). The principle of relativity thus requires a modification of Newton's equation such that for velocities $v \ll c$ Newton's equations remain valid.

6.4.1 Momentum and Energy

We first consider the case of a free particle. The Newton momentum

$$\vec{p} = m_0 \frac{d\vec{r}}{dt} \tag{6.94}$$

is extended to a four-momentum p_μ with components given by

$$p_\mu = m_0 \frac{dx_\mu}{d\tau}, \tag{6.95}$$

where τ is the eigentime of the particle in its rest system and m_0 its restmass. The eigentime τ is related to the time t in the system Σ, where the coordinates x_μ are defined, as follows:

$$t = \gamma\tau; \qquad \gamma = \left(1 - \frac{v^2}{c^2}\right)^{-1/2} = \gamma(v). \tag{6.96}$$

For $v \ll c$ the Lorentz factor $\gamma \to 1$ and the spatial components of (6.95) merge with (6.94). In order to interpret the additional 0th component in (6.95),

$$p_0 = m_0 \frac{dx_0}{d\tau} = \frac{i}{c} m_0 \gamma c^2 \tag{6.97}$$

we recall that the p_μ are components of a four-vector since m_0 and τ are Lorentz scalars. However, the **length** of a four-vector is Lorentz invariant according to Sect. 6.3:

(6.98)

$$\sum_\mu p_\mu^2 = \text{const} = -m_0^2c^2.$$

The right side of (6.98) is obtained as follows: For the spatial components we have

$$p^2 = \sum_{i=1}^3 p_i^2 = m_0^2\gamma^2v^2, \tag{6.99}$$

where v is the magnitude of the velocity $\overrightarrow{v}$ of the particle. Furthermore,

$$p_0^2 = -m_0^2\gamma^2c^2, \tag{6.100}$$

such that

$$\boxed{\sum_\mu p_\mu^2 = m_0^2c^2(\gamma^2\beta^2 - \gamma^2) = -m_0^2c^2.} \tag{6.101}$$

To interpret p_0 we expand $\gamma(v)$ for $v \ll c$ as:

$$m_0c^2\gamma = m_0c^2(1 + \frac{\beta^2}{2}\cdots) = m_0c^2 + \frac{1}{2}m_0v^2 + \cdots \tag{6.102}$$

Since the 2nd term on the right side is the non-relativistic (kinetic) energy of the particle, it is reasonable to interpret

$$\boxed{\epsilon = m_0\gamma(v)c^2 = m(v)c^2} \tag{6.103}$$

as the energy of the free particle; the contribution

$$\boxed{\epsilon_0 = m_0c^2} \tag{6.104}$$

is its **rest energy**. Accordingly

$$\boxed{T = \epsilon - \epsilon_0} \tag{6.105}$$

is its relativistic **kinetic energy**. Equation (6.101) can then be written as a relativistic energy-momentum relation:

$$\boxed{\epsilon^2 = c^2(p^2 + m_0^2c^2)} \tag{6.106}$$

and the four-vector (6.95) has the components

$$\boxed{(\tfrac{i}{c}\epsilon, p_1, p_2, p_3).} \tag{6.107}$$

The **equivalence of energy and mass** in (6.103) has been confirmed by a variety of experiments. Some representative examples are:

(1.) **Binding energies of atoms and nuclei**
For the deuteron the mass difference

$$\Delta m = m_p + m_n - m_d \approx 3.5 \cdot 10^{-27} g \tag{6.108}$$

corresponds to an energy

$$\epsilon_d = \Delta m \ c^2 \approx 2.2 \mathrm{MeV}, \tag{6.109}$$

which is the binding energy of the deuteron. In atoms the binding energy is orders of magnitude lower: from

$$m_p + m_e - m_H \approx 2.4 \cdot 10^{-32} g \tag{6.110}$$

it follows for the binding energy of hydrogen

$$\epsilon_H \approx 13.5 \mathrm{eV}. \tag{6.111}$$

(2.) **Energy production in stars**
One of the essential reactions for energy production in stars is the **fusion** of hydrogen (H) to helium (^4He). This elementary process has the mass balance

$$4m_p + 2m_e - m_{^4He} \approx 0.5 \cdot 10^{-25} g, \tag{6.112}$$

which gives about 25 MeV of energy gain.

(3.) **Pair creation and destruction**

When electrons collide with positrons, high-energy γ quanta (hard photons) are produced,

$$e^+ + e^- \to 2\gamma, \tag{6.113}$$

where the energy-momentum balance requires the appearance of 2 γ-quanta. On the other hand, a γ-quantum (> 1.02 MeV $\approx 2m_e c^2$) can be converted into an electron-positron pair,

$$\gamma \to e^+ + e^-, \tag{6.114}$$

if another particle (e.g. an atomic nucleus) ensures the momentum balance.

We generalize **Newton's inertial law** for a free particle,

$$\overrightarrow{p} = \text{const}, \tag{6.115}$$

to

$$\boxed{p_\mu = \text{const}; \qquad \mu = 0, 1, 2, 3,} \tag{6.116}$$

thus also demand that the 0th component, the energy ϵ, is constant.

The generalization (6.116) of (6.115) follows necessarily from the transformation property of p_μ. Since they are components of a four-vector the following holds for a Lorentz transformation (in $x-$ direction with velocity v):

$$\epsilon = \gamma(v)(\epsilon' + vp'_x); p_x = \gamma(v)(p'_x + \tfrac{v}{c^2}\epsilon'); p_y = p'_y; p_z = p'_z. \tag{6.117}$$

The mixing of space and time components leads to the fact that the conservation of momentum and energy are only possible simultaneously!

We define the **rest system** of a particle (Σ') by:

$$\epsilon' = m_0 c^2; \qquad p'_x = p'_y = p'_z = 0, \tag{6.118}$$

such that—according to (6.117)—in another inertial system Σ we get:

$$\epsilon = \gamma(v)\epsilon' = m_0\gamma c^2 = m(v)c^2; \qquad p_x = \gamma(v)\tfrac{v}{c^2}\epsilon' = m(v)v; \qquad p_y = p_z = 0. \tag{6.119}$$

Note:

For particles with $m_0 = 0$ such as photons, a rest system cannot be defined because according to (6.118), (6.119) in every inertial system we would obtain $p_\mu = 0$, $\mu = 0, 1, 2, 3$.

6.4.2 Scattering Problems

For the relativistic description of collision processes we define energy and momentum for N particles as:

$$\overrightarrow{P} = \sum_{i=1}^{N} \overrightarrow{p_i}; \qquad \epsilon = \sum_{i=1}^{N} \epsilon_i, \tag{6.120}$$

where $\overrightarrow{p_i}$ are the spatial components of the momentum of particle i, ϵ_i its energy according to (6.103).

We now consider the collision of two particles

$$1 + 2 \rightarrow 3 + 4, \tag{6.121}$$

where (1, 2) denote the particles before the collision and (3, 4) after the collision. Since asymptotically there are free particles (before and after the collision), the conservation of momentum must apply:

$$\overrightarrow{p_1} + \overrightarrow{p_2} - \overrightarrow{p_3} - \overrightarrow{p_4} = 0. \tag{6.122}$$

But if the 3 spatial components of a four-vector disappear, then the 0th component must also disappear according to (6.117),

(6.123)

$$\epsilon_1 + \epsilon_2 - \epsilon_3 - \epsilon_4 = 0,$$

i.e. energy conservation must also hold such that the conservation of momentum holds in every inertial system. Energy and momentum conservation—as Lorentz-invariant statements—can only hold simultaneously as pointed out above!

Example: Compton effect

We investigate the scattering of a photon from a free, initially resting electron. The energy of the photon depends on the frequency ω of the light wave according to

$$\epsilon_\gamma = \hbar\omega, \tag{6.124}$$

where $\hbar \approx 197$ MeV fm/c is Planck's constant. It follows for the momentum (6.106)

$$p_\gamma = \frac{\epsilon_\gamma}{c} = \frac{\hbar\omega}{c} = \frac{2\pi\hbar}{\lambda} = \hbar k, \tag{6.125}$$

since the photon has no rest mass (Fig. 6.4).

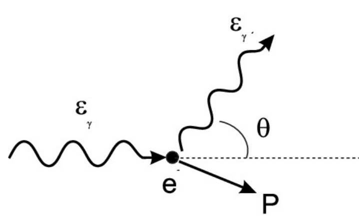

Fig. 6.4 Scattering of a photon on a free, initially resting electron. The final momentum of the electron is denoted by $\vec{P}$ while the scattering angle of the photon is θ

According to energy and momentum conservation:

$$\vec{P} = \hbar(\vec{k} - \vec{k}') \tag{6.126}$$

and

$$c\sqrt{P^2 + m_0^2c^2} - m_0c^2 = \hbar(\omega - \omega') \tag{6.127}$$

for the kinetic energy of the electron after the collision. We square both equations

$$P^2 = \hbar^2(k^2 - 2kk' \cos\theta + k'^2), \tag{6.128}$$

as well as

$$P^2 + m_0^2c^2 = m_0^2c^2 + \hbar^2(k - k')^2 + 2m_0\hbar c(k - k'), \tag{6.129}$$

and consider the difference:

(6.130)

$$\left(\frac{1}{k'} - \frac{1}{k}\right) = \frac{\hbar}{m_0 c}(1 - \cos\theta).$$

We get the change in the wave number of light as a function of the scattering angle θ. The experimental confirmation of (6.130) is an important support for the description of a light wave by **photons**, massless particles, whose energy and momentum are defined by (6.124), (6.125).

6.4.3 Equations of Motion

In generalizing Newton's definition of force we introduce a **four-force in Minkowski space** via its components as:

$$\mathscr{F}_\mu = \frac{dp_\mu}{d\tau} = \gamma(v)\frac{dp_\mu}{dt}. \tag{6.131}$$

Here $d\tau$ is defined in the **instantaneous** rest frame of the particle as a **differential** proper time. The spatial components of (6.131) result in the relativistic generalization of Newton's equations of motion:

$$\frac{d\overrightarrow{p}}{dt} = \gamma^{-1}\mathscr{F} = \overrightarrow{F}, \tag{6.132}$$

where $\overrightarrow{F}$ e.g. stands for the Lorentz force. With (6.99) we can also write:

$$\frac{d}{dt}(m_0\gamma(v)\overrightarrow{v}) = \overrightarrow{F}, \tag{6.133}$$

which for $v \ll c, \gamma \to 1$ leads to the non-relativistic equation of motion:

$$m_0\frac{d\overrightarrow{v}}{dt} = m_0\overrightarrow{a} = \overrightarrow{F}. \tag{6.134}$$

Equation (6.133) has two possible interpretations:

(i) One keeps the non-relativistic velocity $\overrightarrow{v}$ and accepts a velocity-dependent mass,

$$\frac{d}{dt}(m(v)\overrightarrow{v}) = \overrightarrow{F}, \tag{6.135}$$

with

$$m(v) = \gamma(v)\ m_0, \tag{6.136}$$

or

(ii) one always works with the rest mass m_0, a Lorentz invariant quantity, and modifies the definition of velocity:

$$m_0\frac{d\overrightarrow{u}}{dt} = \overrightarrow{F} \tag{6.137}$$

with the modified velocity

with the modified velocity

$$\vec{u} = \gamma(v)\ \vec{v}. \tag{6.138}$$

The equations (6.135) and (6.136) show that particles of the rest mass $m_0 \neq 0$ cannot reach the velocity $v = c$, since for

$$m(v) \to \infty \tag{6.139}$$

in case of $v \to c$ an infinitely large energy would be necessary.

To discuss the component $\mathscr{F}_0$ we use:

$$\sum_\mu \mathscr{F}_\mu p_\mu = \frac{1}{2}\frac{d}{d\tau}(\sum_\mu p_\mu^2) = 0 \tag{6.140}$$

due to (6.98), which gives

$$\sum_{i=1}^3 \mathscr{F}_i p_i = -\mathscr{F}_0 p_0 \tag{6.141}$$

or according to (6.95), (6.97)

$$\mathscr{F}_0 = \frac{i}{c}\vec{\mathscr{F}}\cdot\vec{v} = \frac{i}{c}\gamma(v)\ \vec{F}\cdot\vec{v}. \tag{6.142}$$

Since $\vec{F}\cdot\vec{v}$ is the work done by the force $\vec{F}$ on the particle per unit of time, we can also write

$$\mathscr{F}_0 = \frac{i}{c}\gamma(v)\frac{d\epsilon}{dt} \tag{6.143}$$

or

$$F_0 = \gamma(v)^{-1}\mathscr{F}_0 = \frac{i}{c}\frac{d\epsilon}{dt} \tag{6.144}$$

as expected according to (6.107). The equations (6.142) and (6.143) confirm once again the **equivalence of energy and mass**.

6.4.4 Lorentz Transformation of the Force

Since $\mathscr{F}_\mu$ are the components of a four-vector, the following holds for a transformation from the current rest system Σ to another inertial system Σ' with the special transformation (6.71):

$$\mathscr{F}_1' = \gamma(v)(\mathscr{F}_1 + i\beta\mathscr{F}_0) = \gamma(v)\mathscr{F}_1; \mathscr{F}_2' = \mathscr{F}_2; \mathscr{F}_3' = \mathscr{F}_3, \tag{6.145}$$

since $\mathscr{F}_0 = 0$ in Σ as the instantaneous rest system according to (6.142). In short:

$$\vec{\mathscr{F}}_\perp{}' = \vec{\mathscr{F}}_\perp; \qquad \vec{\mathscr{F}}_\parallel{}' = \gamma(v)\ \vec{\mathscr{F}}_\parallel. \tag{6.146}$$

Due to (6.132) the inverse relations hold

(6.147)

$$\overrightarrow{F}_{\perp}{}' = \sqrt{1 - v^2/c^2}\ \overrightarrow{F}_{\perp}; \qquad \overrightarrow{F}_{\parallel}{}' = \overrightarrow{F}_{\parallel},$$

since in the current rest system Σ we have

$$\gamma(v) = \gamma(0) = 1. \tag{6.148}$$

Result:

We have extended the basic concepts and basic equations of Newton's mechanics to relativistic mechanics in such a way that

(i) Newtonian mechanics is regained in the limit $v \ll c$,

(ii) the modified basic equations are covariant with respect to Lorentz transformations.

In summarizing this chapter we have introduced Einstein's special theory of relativity and replaced the Galilei transformation between inertial systems in Newtonian dynamics by the Lorentz transformation that keeps the velocity of light c invariant in all inertial systems. To this aim we explicitly have derived the Lorentz transformation (in a single spatial dimension) and discussed its implications: Lorentz contraction, time dilation, simultaneity in moving systems as well as causality and the limiting velocity of signals. Some mathematical aspects of the Lorentz group of transformations have been discussed and Lorentz scalars, four-vectors and Lorentz tensors been identified as well as corresponding physical quantities like four-current densities. We, furthermore, have discussed the relativistic dynamics by introducing the energy-momentum four-vector, which is conserved in all four components for closed systems and discussed scattering problems. As an example the important problem of Compton scattering of a photon on a resting charge q has been computed explicitly. The derivation of the Lorentz transformation of the force has completed this chapter.

W. Cassing, *Theoretical Physics Compact I*
https://doi.org/10.1007/978-3-031-95512-9_7

7. Formal Structure of Mechanics

Wolfgang Cassing[1]
(1) University of Giessen, Gießen, Hessen, Germany

The equations of motion of Newtonian mechanics can be written in different ways—depending on the choice of coordinates—and in principle all independent choices have equal rights. However, some choices facilitate the solutions of the equations of motion and others might cause severe problems. It is thus of general interest to find 'optimal' coordinates for the description, which is also of practical help, if the system is subject to constraints that require the introduction of 'coercive forces', which often are difficult to define. It is thus meaningful to define 'generalized coordinates' that fulfill the constraints and also reduce the complexity of the problem by reducing the number of (linear independent) degrees of freedom. The equations of motion in generalized coordinates then are derived from Newton's equations of motion. It will be found that these equations can also be generated by a variational principle, which specifies a Lagrange function L, which is given by the difference between the kinetic and potential energy in case of conservative forces. An important consequence is that the Lagrange equations of motion can also be applied to other areas of physics. Generalized momenta are defined by the derivative of the Lagrange function with respect to the generalized velocities. Accordingly, if the Lagrange function does not depend on a specific coordinate, e.g. the azimuthal angle φ, the corresponding generalized momentum (here angular momentum) is a constant of motion. This suggests to transform the formulation to phase-space variables given by coordinates and their associated momenta, which is carried out by a Legendre transformation defining the Hamilton function H. In case of conservative forces the latter just gives the energy of the system in phase-space variables. The variational principle thus can be reformulated in terms of Hamilton's (equivalent) variational principle which leads to the canonical equations of motion. The latter will be illustrated for a couple of examples. Furthermore, it will be shown again that—for a closed system—the translational invariance leads to the conservation of the total momentum, the rotational invariance to the conservation of total angular momentum, and the invariance with respect to time translations to the conservation of the total energy.

7.1 Generalized Coordinates

7.1.1 Constraints

The starting point of Newtonian mechanics are the equations of motion for N particles in cartesian coordinates:

$$m_i \ddot{\vec{r}}_i = \vec{F}_i \quad ; \quad i = 1, 2, \ldots, N \, . \tag{7.1}$$

Difficulties arise if the motion of the system is subject to **constraints**:

1. The coordinates $\vec{r}_i$ then are dependent on the constraints.
2. In order to comply with certain constraints one must introduce **coercive forces** which are not

explicitly specified, but in some cases can only be determined from the solution.

Classification of constraints:

1. **Holonome** conditions:
   (a) Scleronome conditions:
      Examples:
      - rigid body

        $$\left(\overrightarrow{r}_i - \overrightarrow{r}_j\right)^2 - c_{ij}^2 = 0 \quad ; \quad i, j = 1, 2, \ldots, N \,. \tag{7.2}$$

      - ball pendulum of length *l*

        $$x^2 + y^2 + z^2 - l^2 = 0. \tag{7.3}$$

   (b) Rheonome conditions

      $$f(\overrightarrow{r}_1, \ldots \overrightarrow{r}_N, t) = 0 \tag{7.4}$$

      contain an explicit time dependence.
      Example: 'Pearl' on a straight rotating wire.
2. **Nonholonomic** conditions
   explicitly require the solution of the equations of motion!
   Example: Gas molecules in a spherical container, $r_i \leq R$.

For **holonome** conditions we can solve the problem by introducing **generalized coordinates** q_j such that for

$$\overrightarrow{r}_i = \overrightarrow{r}_i(q_1, \ldots, q_s, t) \tag{7.5}$$

the *m* constraints

$$\boxed{f_r(\overrightarrow{r}_1, \ldots, \overrightarrow{r}_N, t) = 0 \quad ; \quad r = 1, 2, \ldots, m} \tag{7.6}$$

are identically fulfilled in the new variables q_j and *t*. The variables q_j are independent of each other; if *m* constraints are given, then for *N* particles we have

$$s = 3N - m \leq 3N \tag{7.7}$$

generalized coordinates q_j.

7.1.2 Equations of Motion in Generalized Coordinates

Starting from Newton's equations of motion we form the following $(3N-m)$ differential equations (with $\overrightarrow{r_i}=\overrightarrow{r_i}(q_l)$):

$$\sum_{i=1}^N m_i\ddot{\overrightarrow{r_i}}\cdot\frac{\partial\overrightarrow{r_i}}{\partial q_l}=\sum_{i=1}^N \overrightarrow{F_i}\cdot\frac{\partial\overrightarrow{r_i}}{\partial q_l}=Q_l\ . \tag{7.8}$$

We write the left side as:

$$m_i\ddot{\overrightarrow{r_i}}\cdot\frac{\partial\overrightarrow{r_i}}{\partial q_l}=\frac{d}{dt}\left(m_i\dot{\overrightarrow{r_i}}\cdot\frac{\partial\overrightarrow{r_i}}{\partial q_l}\right)-m_i\dot{\overrightarrow{r_i}}\cdot\frac{d}{dt}\left(\frac{\partial\overrightarrow{r_i}}{\partial q_l}\right), \tag{7.9}$$

using

$$m_i\dot{\overrightarrow{r_i}}\cdot\frac{\partial\overrightarrow{r_i}}{\partial q_l}=m_i\overrightarrow{v_i}\cdot\frac{\partial\overrightarrow{v_i}}{\partial\dot{q}_l}=\frac{\partial}{\partial\dot{q}_l}\left(\frac{1}{2}m_iv_i^2\right)=\frac{\partial}{\partial\dot{q}_l}T_i\quad, \tag{7.10}$$

because

$$\frac{\partial}{\partial\dot{q}_l}\overrightarrow{v_i}=\frac{\partial}{\partial\dot{q}_l}\dot{\overrightarrow{r_i}}=\frac{\partial}{\partial\dot{q}_l}\left[\sum_j\frac{\partial\overrightarrow{r_i}}{\partial q_j}\dot{q}_j+\frac{\partial\overrightarrow{r_i}}{\partial t}\right]=\frac{\partial\overrightarrow{r_i}}{\partial q_l}\ , \tag{7.11}$$

since $\partial\overrightarrow{r_i}/\partial q_j$ and $\partial\overrightarrow{r_i}/\partial t$ do not depend on $\dot{q}_j$. With

$$\frac{d}{dt}\left(\frac{\partial\overrightarrow{r_i}}{\partial q_l}\right)=\sum_j\frac{\partial^2\overrightarrow{r_i}}{\partial q_j\partial q_l}\dot{q}_j+\frac{\partial^2\overrightarrow{r_i}}{\partial q_l\partial t}=\frac{\partial}{\partial q_l}\left(\sum_j\frac{\partial\overrightarrow{r_i}}{\partial q_j}\dot{q}_j+\frac{\partial\overrightarrow{r_i}}{\partial t}\right)=\frac{\partial}{\partial q_l}\overrightarrow{v_i} \tag{7.12}$$

we get for the 2nd term on the right side of (7.9):

$$m_i\overrightarrow{v_i}\cdot\frac{d}{dt}\left(\frac{\partial\overrightarrow{r_i}}{\partial q_l}\right)=m_i\overrightarrow{v_i}\cdot\frac{\partial}{\partial q_l}\overrightarrow{v_i}=\frac{\partial}{\partial q_l}\left(\frac{1}{2}m_iv_i^2\right)=\frac{\partial}{\partial q_l}T_i. \tag{7.13}$$

With the kinetic energy

$$\boxed{T=\sum_{i=1}^N T_i=\frac{1}{2}\sum_{i=1}^N m_iv_i^2=T(q_j,\dot{q}_j,t)} \tag{7.14}$$

we obtain after summation over all particles *i*:

$$\boxed{\frac{d}{dt}\left(\frac{\partial T}{\partial\dot{q}_l}\right)-\frac{\partial}{\partial q_l}T=Q_l\qquad .} \tag{7.15}$$

To interpret the quantities Q_l it is sufficient to consider the case, where the time *t* does not occur explicitly. Then the work carried out by the forces $\overrightarrow{F_i}$ for infinitesimal displacements $d\overrightarrow{r_i}$ of the particles, which comply with the constraints, is given by:

$$dW=\sum_{i=1}^N\overrightarrow{F_i}\cdot d\overrightarrow{r_i}=\sum_{i=1}^N\sum_{l=1}^s\overrightarrow{F_i}\cdot\frac{\partial\overrightarrow{r_i}}{\partial q_l}dq_l=\sum_{l=1}^s Q_ldq_l. \tag{7.16}$$

This suggests that the quantities Q_l can be considered as **generalized forces**. Since the displacements dq_l were introduced in such a way that the constraints are fulfilled, coercive forces cannot contribute, since they do only serve to comply with the mandatory conditions. This implies that when calculating the Q_l from the forces $\overrightarrow{F_i}$ (which additionally contain the constraining forces) the coercive forces cancel out.

7.1.3 Conservative Forces

We consider the case, where a function $U = U(q_j) \neq U(\dot{q}_j)$ exists such that

$$\boxed{Q_l = -\frac{\partial U}{\partial q_l}. \tag{7.17}}$$

We then define the **Lagrange function** of the system by

$$\boxed{L = T - U \ , \tag{7.18}}$$

and from (7.15) get the **Lagrange equation of the second kind**

$$\boxed{\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_l}\right) = \frac{\partial L}{\partial q_l} \ . \tag{7.19}}$$

In analogy to Newton's equations

$$\boxed{p_l = \frac{\partial L}{\partial \dot{q}_l} \tag{7.20}}$$

defines **generalized momenta**. Then (for $T = T(\dot{q}_i)$)

$$\boxed{\frac{d}{dt}p_l = \dot{p}_l = \frac{\partial L}{\partial q_l} = -\frac{\partial U}{\partial q_l} = Q_l \tag{7.21}}$$

achieves the form of a Newtonian equation of motion.

7.1.4 Examples

1. **Particles without constraints**: In this case the generalized coordinates are $q_l = (x, y, z)$; we form

$$T = \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \quad , \tag{7.22}$$

and get

$$\frac{\partial T}{\partial x} = \frac{\partial T}{\partial y} = \frac{\partial T}{\partial z} = 0 \quad ; \quad \frac{\partial T}{\partial \dot{x}} = m\dot{x}; \quad \frac{\partial T}{\partial \dot{y}} = m\dot{y}; \quad \frac{\partial T}{\partial \dot{z}} = m\dot{z}. \tag{7.23}$$

With

$$Q_x = F_x \tag{7.24}$$

we obtain

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{x}}\right) = m\ddot{x} = F_x = Q_x \tag{7.25}$$

etc. for y, z, which just gives Newton's equations of motion.

2. **Motion of a particle in the plane**: For convenience we use polar coordinates (see Fig. 7.1), i.e.

$$x = r\cos\varphi \quad ; \quad y = r\sin\varphi \ . \tag{7.26}$$

Then we get for the velocities:

$$\dot{x} = \dot{r}\frac{\partial x}{\partial r} + \dot{\varphi}\frac{\partial x}{\partial \varphi} = \dot{r}\cos\varphi - r\dot{\varphi}\sin\varphi \quad ; \quad \dot{y} = \dot{r}\sin\varphi + r\dot{\varphi}\cos\varphi \quad . \tag{7.27}$$

The kinetic energy amounts to:

$$T = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) = \frac{m}{2}(\dot{r}^2 + r^2\dot{\varphi}^2) \tag{7.28}$$

and

$$\frac{\partial T}{\partial r} = mr\dot{\varphi}^2 \quad ; \quad \frac{\partial T}{\partial \varphi} = 0 \tag{7.29}$$

$$\frac{\partial T}{\partial \dot{r}} = m\dot{r} \quad ; \quad \frac{\partial T}{\partial \dot{\varphi}} = mr^2\dot{\varphi}. \tag{7.30}$$

For the forces we obtain

$$Q_r = \overrightarrow{F}\cdot\frac{\partial \overrightarrow{r}}{\partial r} = \overrightarrow{F}\cdot\frac{\overrightarrow{r}}{r} = \overrightarrow{F}\cdot\overrightarrow{e}_r = F_r \tag{7.31}$$

$$Q_\varphi = \overrightarrow{F}\cdot\frac{\partial \overrightarrow{r}}{\partial \varphi} = r\overrightarrow{F}\cdot\overrightarrow{e}_\varphi = rF_\varphi. \tag{7.32}$$

The Lagrange equations then are:

$$m\ddot{r} - mr\dot{\varphi}^2 = F_r \quad ; \quad \frac{d}{dt}(mr^2\dot{\varphi}) = rF_\varphi \ . \tag{7.33}$$

In the right equation, $mr^2\dot{\varphi}$ is the **angular momentum**, whose temporal change is given by the **torque** $rF_\varphi = Q_\varphi$, which plays the role of a generalized force. **Special case**: For the flat pendulum (see Fig. 7.2) we have the constraint:

$$r - l = 0 \quad , \tag{7.34}$$

if l is the constant length of the pendulum. In this case T reduces to

$$T = \frac{ml^2}{2}\dot{\varphi}^2 \quad ; \quad \frac{\partial T}{\partial \dot{\varphi}} = ml^2\dot{\varphi}. \tag{7.35}$$

The Lagrange equations with $U(\varphi) = mgl(1 - \cos\varphi)$ simplify, too:

$$ml\ddot{\varphi} = F_\varphi = -mg \sin \varphi, \tag{7.36}$$

$$\ddot{\varphi} + \omega_0^2 \sin \varphi = 0 \quad \text{with} \quad \omega_0^2 = \tfrac{g}{l}. \tag{7.37}$$

For small deflections we may approximate $\sin \varphi \approx \varphi$ and thus get

$$\ddot{\varphi} + \omega_0^2 \varphi = 0. \tag{7.38}$$

3. **Atwood's machine**
   The constraint (see Fig. 7.3)

$$x_1 + x_2 = l = x + (l - x) \tag{7.39}$$

is identically fulfilled in the coordinate $q = x$. Then the kinetic energy is given by:

$$T = \tfrac{1}{2}(m_1 + m_2)\dot{x}^2 \quad , \tag{7.40}$$

and the potential energy by

$$U = -m_1 g x - m_2 g(l - x). \tag{7.41}$$

The Lagrangian is

$$L = \tfrac{m_1+m_2}{2}\dot{x}^2 + m_1 g x + m_2 g(l - x) \tag{7.42}$$

and the Lagrange equation reads

$$\tfrac{d}{dt}\left(\tfrac{\partial L}{\partial \dot{x}}\right) = (m_1 + m_2)\ddot{x} \tag{7.43}$$

We get:

$$(m_1 + m_2)\ddot{x} = (m_1 - m_2)g. \tag{7.44}$$

4. **Pearl on a rotating wire**
   For (see Fig. 7.4)

$$x = r \cos(\omega t) \quad ; \quad y = r \sin(\omega t) \tag{7.45}$$

the kinetic energy according to (7.28) is given by

$$T = \tfrac{m}{2}(\dot{r}^2 + r^2\omega^2). \tag{7.46}$$

The equations of motion for the force-free case $L = T$ are

$$m\ddot{r} - mr\omega^2 = 0 \tag{7.47}$$

with $mr\omega^2$ as the well known **centrifugal force**.

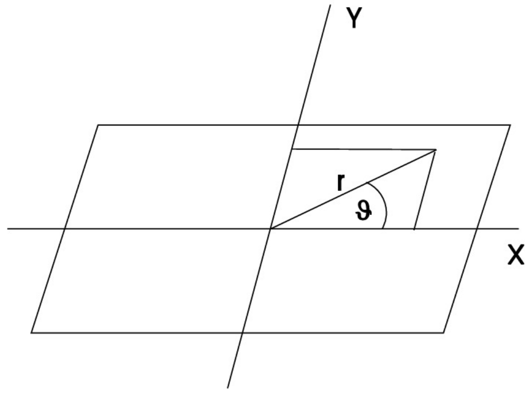

Fig. 7.1 The flat pendulum of length *l*

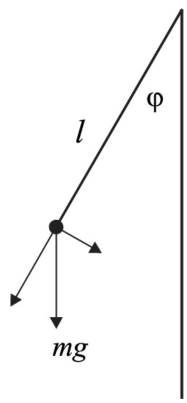

Fig. 7.2 Illustration of Atwood's machine with a rope of length *l*

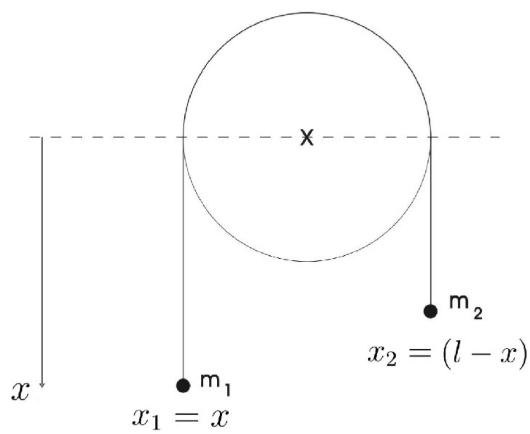

Fig. 7.3 Pearl on a rotating wire

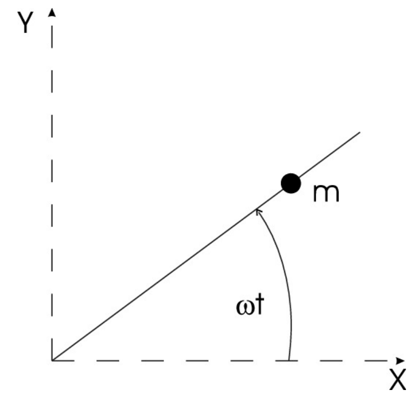

Fig. 7.4 Choice of coordinates for motions in a plane

7.1.5 Velocity-Dependent Forces

The Lagrange equations of second kind hold for velocity-dependent forces, if a function $U(q_i, \dot{q}_i, t)$ exists such that:

$$Q_l = -\frac{\partial U}{\partial q_l} + \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{q}_l}\right). \tag{7.48}$$

An important example for such a velocity-dependent force is the **Lorentz force**, which arises from

$$U(\overrightarrow{r}, \overrightarrow{v}, t) = e(\Phi(\overrightarrow{r}, t) - \overrightarrow{v} \cdot \overrightarrow{A}(\overrightarrow{r}, t)) \tag{7.49}$$

for a particle of charge *e* with velocity $\overrightarrow{v}$ (see electrodynamics). Here $\Phi(\overrightarrow{r}, t)$ is a scalar function (related to the charge density) and $A(\overrightarrow{r}, t)$ a vector field (related to the charge current density). The force (in cartesian coordinates $q_l = (x, y, z)$) is given by:

$$F_x = e(-\operatorname{grad}\Phi - \frac{\partial \overrightarrow{A}}{\partial t} + (\overrightarrow{v} \times \operatorname{rot}\overrightarrow{\mathrm{A}}))_{\mathrm{x}} \quad ,$$

$$F_y = e(-\operatorname{grad}\Phi - \frac{\partial \overrightarrow{A}}{\partial t} + (\overrightarrow{v} \times \operatorname{rot}\overrightarrow{\mathrm{A}}))_{\mathrm{y}} \quad ,$$

$$F_z = e(-\operatorname{grad}\Phi - \frac{\partial \overrightarrow{A}}{\partial t} + (\overrightarrow{v} \times \operatorname{rot}\overrightarrow{A}))_z \quad , \tag{7.50}$$

using

$$\frac{dA_x}{dt} = \frac{\partial A_x}{\partial x} v_x + \frac{\partial A_x}{\partial y} v_y + \frac{\partial A_x}{\partial z} v_z + \frac{\partial A_x}{\partial t} \ . \tag{7.51}$$

Important note: For a **gauge transformation**

$$\overrightarrow{A} \to \overrightarrow{A} + \operatorname{grad}\chi \qquad \Phi \to \Phi - \frac{\partial \chi}{\partial \mathrm{t}} \ , \tag{7.52}$$

where the function $\chi = \chi(\vec{r}, t)$ is arbitrary but continuously differentiable in all variables, the potential transforms as

$$U \to U - e\left(\vec{v} \cdot \operatorname{grad} \chi + \frac{\partial \chi}{\partial t}\right) = U - e\frac{d\chi}{dt} \ . \tag{7.53}$$

The equations of motion then do **not** change for a gauge transformation

$$L \to L' = L + \frac{dg}{dt} \tag{7.54}$$

with any twice continuously differentiable function $g = g(q_l, t)$. Due to

$$\frac{dg}{dt} = \sum_j \frac{\partial g}{\partial q_j}\dot{q}_j + \frac{\partial g}{\partial t} \tag{7.55}$$

we have

$$\frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_l}\left(\frac{dg}{dt}\right)\right) = \frac{d}{dt}\left(\frac{\partial g}{\partial q_l}\right) \quad , \tag{7.56}$$

such that the additional term (7.56) in the Lagrange equation

$$\frac{d}{dt}\left(\frac{\partial g}{\partial q_l}\right) \tag{7.57}$$

is cancelled again by the additional term in the partial derivative with respect to q_l

$$\frac{\partial}{\partial q_l}\left(\frac{dg}{dt}\right), \tag{7.58}$$

since $g = g(q_l, t)$ was assumed to be twice continuously differentiable.

The invariance of the equations of motion for the transformation (7.52) implies that the Lagrange function L itself is not uniquely determined.

We will exploit this property in **field theory** to derive the Lorentz force itself from 'simple considerations'.

7.2 Hamilton's Variational Principle

7.2.1 Variational Principle and Euler's Equations

Let a system of N particles with m holonome constraints be described by generalized coordinates q_i. The values of the coordinates at a fixed time t then determine a **point** in the **configuration space** with dimension $s = 3N - m$ that is spanned by the coordinates q_i. The temporal evolution of the system corresponds to a **trajectory in configuration space** with the time t as a parameter of the trajectory.

The **actual** trajectory—traversed by the system—is the solution of s Lagrange equations (7.19). It is uniquely determined if—for fixing the $2s$ integration constants—

1. at a time t_1 apart from the $q_i(t_1)$ also the generalized velocities $\dot{q}_i(t_1)$ are known, **or**
2. the trajectory points $q_i(t_1)$ and $q_i(t_2)$ are given for different times $t_1 \neq t_2$.

In the latter case we can indeed characterize the actual trajectory relative to neighboring trajectories, which also pass through the points $q_i(t_1)$ and $q_i(t_2)$, by the fact that the action

$$\boxed{S[q_i, \dot{q}_i] = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t)\ dt,} \tag{7.59}$$

given by the time integral of the Lagrange function $L(q_i, \dot{q}_i, t)$, has an extremum for the actual trajectory, i.e.

$$\delta S[q_i, \dot{q}_i] = 0. \tag{7.60}$$

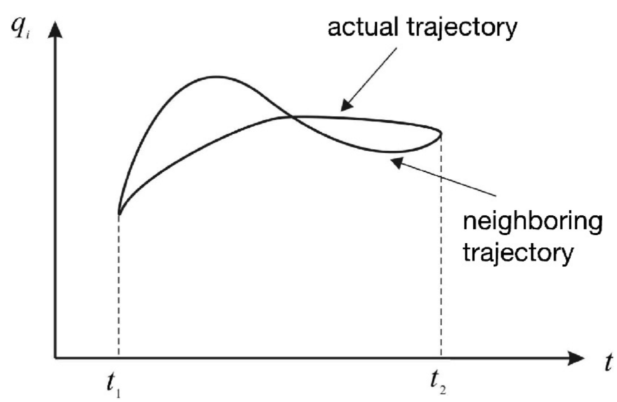

Fig. 7.5 Illustration of an actual trajectory and a neighboring trajectory, which pass through the same points at t_1 and t_2

In order to explain the variation principle (7.59) in more detail, we consider any neighboring trajectory to the actual trajectory $q_i(t)$ (for small ε),

$$q_i'(t) = q_i(t) + \varepsilon\eta_i(t), \tag{7.61}$$

with the property that $q_i'(t)$ matches with the trajectory $q_i(t)$ at the times t_1 and t_2 (see Fig. 7.5), i.e.

$$\eta_i(t_1) = \eta_i(t_2) = 0. \tag{7.62}$$

Then for

$$\tilde{S}(\varepsilon) = \int_{t_1}^{t_2} L(q_i + \varepsilon\eta_i, \dot{q}_i + \varepsilon\dot{\eta}_i, t)\ dt \tag{7.63}$$

(after applying (7.60)) must hold:

$$\left(\frac{\partial \tilde{S}}{\partial \varepsilon}\right)_{\varepsilon=0} = 0. \tag{7.64}$$

Explicitly this leads to:

$$\int_{t_1}^{t_2} \sum_i \left\{ \frac{\partial L}{\partial q_i} \eta_i + \frac{\partial L}{\partial \dot{q}_i} \dot{\eta}_i \right\} dt = 0. \tag{7.65}$$

By partial integration in time for the 2nd term

$$\int_{t_1}^{t_2} \frac{\partial L}{\partial \dot{q}_i} \dot{\eta}_i dt = \left[\frac{\partial L}{\partial \dot{q}_i} \eta_i \right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \eta_i \ dt \tag{7.66}$$

it follows, since the integrated term [...] vanishes according to the assumption (7.62), that (7.65) becomes

$$\int_{t_1}^{t_2} \sum_i \left\{ \frac{\partial L}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \right\} \eta_i \ dt = 0. \tag{7.67}$$

Since the functions $\eta_i(t)$ are linearly independent and arbitrary in the time interval $t_1 < t < t_2$, the **Euler equations of the variation principle** are identical to the **Lagrange equations**

$$\boxed{\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0.} \tag{7.68}$$

Note: The variation principle not only offers an elegant formulation, which is equivalent to the equations of motion of classical non-relativistic mechanics, but can also be applied to other areas of physics, such as elastic media, electrodynamics, and field theory of elementary particles.

7.2.2 Canonical Equations

For the transition from classical mechanics to **quantum mechanics** and for **statistical mechanics** it will be useful to transform from the variables $\{q_i, \dot{q}_i\}$ to an equivalent set of variables $\{q_i, p_i\}$. In the following we want to derive **canonical equations** in the variables q_i, p_i which are equivalent to those formulated in the variables q_i, $\dot{q}_i$ (7.68). Instead of the Lagrange function $L = L(q_i, \dot{q}_i, t)$ a new function $H = H(q_i, p_i, t)$, the **Hamilton function** of the system, will be introduced.

The transition in the variables

$$\{q_i, \dot{q}_i, t\} \rightarrow \{q_i, p_i, t\} \tag{7.69}$$

as well as

$$L(q_i, \dot{q}_i, t) \rightarrow H(q_i, p_i, t) \tag{7.70}$$

is performed by a **Legendre transformation** .

To explain the Legendre transformation, we first consider—as a simple example—a function $f(x, y)$ of the independent variables x, y. Then the total differential of f can be written as:

$$df = vdx + udy \tag{7.71}$$

with

$$v = \frac{\partial f}{\partial x} \qquad u = \frac{\partial f}{\partial y}, \tag{7.72}$$

where $v(x, y)$ and $u(x, y)$ are connected via

$$\frac{\partial v}{\partial y} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial u}{\partial x}, \tag{7.73}$$

if f is assumed to be twice continuously differentiable. Now in the transformation

$$\{x, y\} \to \{x, u\}, \tag{7.74}$$

the function

$$uy - f(x, y) = g(x, u) \tag{7.75}$$

can be represented alone by the independent variables (x, u).

Proof: For the total differential of g, which according to (7.75) is a function of x, y, u at first sight, we get:

$$dg = udy + ydu - df = udy + ydu - vdx - udy = -vdx + ydu = \frac{\partial g}{\partial x}dx + \frac{\partial g}{\partial u}du, \tag{7.76}$$

i.e. the function g in fact depends only on x and $u = \partial f/\partial y$ and no longer on y (q.e.d.). After comparing the coefficients we find:

$$v = -\frac{\partial g}{\partial x} = \frac{\partial f}{\partial x}; \qquad y = \frac{\partial g}{\partial u}. \tag{7.77}$$

In analogy we now introduce the **Hamilton function** H by:

$$\boxed{H(q_i, p_i, t) = \sum_{i=1}^{s} \dot{q}_i p_i - L(q_i, \dot{q}_i, t).} \tag{7.78}$$

Forming the total differential of H according to the definition (7.78),

$$dH = \sum_{i=1}^{s}\left\{\dot{q}_i dp_i + p_i d\dot{q}_i - \frac{\partial L}{\partial q_i}dq_i - \frac{\partial L}{\partial \dot{q}_i}d\dot{q}_i\right\} - \frac{\partial L}{\partial t}dt, \tag{7.79}$$

we obtain with the definition of

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{7.80}$$

for the total differential of H:

$$dH = \sum_{i=1}^{s} \dot{q}_i dp_i - \sum_{i=1}^{s} \frac{\partial L}{\partial q_i}dq_i - \frac{\partial L}{\partial t}dt. \tag{7.81}$$

The comparison with

$$dH = \sum_{i=1}^{s} \frac{\partial H}{\partial q_i}dq_i + \sum_{i=1}^{s} \frac{\partial H}{\partial p_i}dp_i + \frac{\partial H}{\partial t}dt \tag{7.82}$$

shows that (using the Lagrange equation):

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \quad ; \quad \dot{p}_i = -\frac{\partial H}{\partial q_i} \tag{7.83}$$

and

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \tag{7.84}$$

The $2s$ differential equations of first order (7.83), which are denoted as **canonical differential equations**, replace the s differential equations of second order (7.68).

The system at time t is now represented by a **point** in **phase space** with the dimension $2s$, which is spanned by the independent variables (q_i, p_i). Contrary to the configuration space, where there is an infinite manifold of orbits at q_i, that are distinguished by the generalized velocities $\dot{q}_i$, there is only a single **trajectory** in phase space through each point (q_i, p_i) (in case of given forces), since the values of q_i and p_i at a fixed point in time uniquely determine the temporal evolution of the system.

Note: The canonical equations (7.83) can also be derived from Hamilton's variational principle, i.e.:

$$\int_{t_1}^{t_2}\{\sum_i p_i\dot{q}_i - H(q_i,p_i,t)\}dt = \text{extremum} \tag{7.85}$$

Equation (7.85) is equivalent to (7.60), since (7.85) arises from (7.60) using (7.78). The 'variations' of q_i and p_i are considered to be **independent** from each other.

7.2.3 Examples

(1.) For the **one-dimensional harmonic oscillator** the Lagrangian reads

$$L = \frac{1}{2}mv^2 - \frac{D}{2}x^2, \qquad p = \frac{\partial L}{\partial v} = m\dot{q} = mv, \tag{7.86}$$

accordingly the Hamiltonian is:

$$H = \dot{q}p - L = 2T - T + U = T + U = \frac{1}{2m}(p^2 + \omega_0^2 m^2 x^2) \tag{7.87}$$

with $\omega_0^2 = D/m$. The canonical differential equations then are given by:

$$\dot{q}_i = \dot{x} = \frac{\partial H}{\partial p} = \frac{p}{m} = v \tag{7.88}$$

and

$$\dot{p} = -\frac{\partial H}{\partial x} = -Dx, \tag{7.89}$$

or together (with $\omega_0^2 = D/m$):

$$\ddot{x} + \omega_0^2 x = 0. \tag{7.90}$$

(2.) For a **particle in the electromagnetic field** with charge e and velocity $\vec{v}$ the Lagrangian is

$$L = T - e\Phi + e\vec{v}\cdot\vec{A} \tag{7.91}$$

with a scalar potential Φ and a vector potential $\vec{A}$. The momentum components are:

$$p_x = \frac{\partial L}{\partial v_x} = mv_x + eA_x,\ \ p_y = \frac{\partial L}{\partial v_y} = mv_y + eA_y,\ \ p_z = \frac{\partial L}{\partial v_z} = mv_z + eA_z \tag{7.92}$$

and with $m\vec{v} = \vec{p} - e\vec{A}$ we get

$$H = \dot{\vec{r}} \cdot \vec{p} - L = \vec{v} \cdot \vec{p} - T + e\Phi - e\vec{v} \cdot \vec{A}$$

$$= \vec{v} \cdot (m\vec{v} + e\vec{A}) - T + e\Phi - e\vec{v} \cdot \vec{A}$$

$$= mv^2 - T + e\Phi = T + e\Phi = \frac{1}{2m}\left(\vec{p} - e\vec{A}\right)^2 + e\Phi. \tag{7.93}$$

The canonical differential equations e.g. for the components in *x* direction read:

$$\dot{x} = \frac{\partial H}{\partial p_x} = \frac{2mv_x}{2m} = v_x = \frac{1}{m}(p_x - eA_x) \tag{7.94}$$

$$\dot{p}_x = -\frac{\partial H}{\partial x} = -e\frac{\partial \Phi}{\partial x} + \frac{e}{m}\left(\vec{p} - e\vec{A}\right) \cdot \frac{\partial \vec{A}}{\partial x}. \tag{7.95}$$

In summary:

$$m\ddot{x} = -e\frac{\partial \Phi}{\partial x} + \frac{e}{m}(\vec{p} - e\vec{A}) \cdot \frac{\partial \vec{A}}{\partial x} - e\frac{dA_x}{dt}, \tag{7.96}$$

or

$$m\ddot{x} = -e\left(\frac{\partial \Phi}{\partial x} + \frac{\partial A_x}{\partial t}\right) + e\left(\vec{v} \times \left(\vec{\nabla} \times \vec{A}\right)\right)_x. \tag{7.97}$$

(3.) **Rotating coordinate systems**
The relationship between two coordinate systems rotating around the *z*-axis with relative angular velocity $\vec{\omega} = (0, 0, \omega) = \omega\vec{e}_z$ is:

$$x = x' \cos \omega t - y' \sin \omega t$$

$$y = x' \sin \omega t + y' \cos \omega t$$

$$z = z'. \tag{7.98}$$

For the time derivatives we get:

$$\dot{x} = \dot{x}' \cos \omega t - x'\omega \sin \omega t - \dot{y}' \sin \omega t - y'\omega \cos \omega t$$

$$\dot{y} = \dot{x}' \sin \omega t + x'\omega \cos \omega t + \dot{y}' \cos \omega t - y'\omega \sin \omega t$$

$$\dot{z} = \dot{z}'. \tag{7.99}$$

This leads to the kinetic energy:

$$T = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) = \frac{m}{2}\left(\dot{x}'^2 + \dot{y}'^2 + \dot{z}'^2\right) + m\omega(x'\dot{y}' - \dot{x}'y') + \frac{m\omega^2}{2}\left(x'^2 + y'^2\right). \tag{7.100}$$

For velocity-independent potentials $U(x, y, z)$ the momentum components then read:

$$p'_x = \frac{\partial L}{\partial \dot{x}'} = \frac{\partial T}{\partial \dot{x}'} = m(\dot{x}' - \omega y') \tag{7.101}$$

$$p'_y = \frac{\partial L}{\partial \dot{y}'} = \frac{\partial T}{\partial \dot{y}'} = m(\dot{y}' + \omega x') \tag{7.102}$$

$$p'_z = \frac{\partial L}{\partial \dot{z}'} = \frac{\partial T}{\partial \dot{z}'} = m\dot{z}'. \tag{7.103}$$

This results in the Hamilton function in the momenta p'_x, p'_y, p'_z and coordinates x', y', z':

$$H = \overrightarrow{p'} \cdot \overrightarrow{v'} - T + U = \frac{1}{2\,m}({p'}^2_x + {p'}^2_y + {p'}^2_z) + \omega(p'_x y' - p'_y x') + U. \tag{7.104}$$

The canonical equations are:

$$\dot{x}' = \frac{\partial H}{\partial p'_x} = \frac{p'_x}{m} + \omega y' = v'_x \tag{7.105}$$

$$\dot{y}' = \frac{\partial H}{\partial p'_y} = \frac{p'_y}{m} - \omega x' = v'_y \tag{7.106}$$

$$\dot{z}' = \frac{\partial H}{\partial p'_z} = \frac{p'_z}{m} = v'_z. \tag{7.107}$$

As in the 2nd example, $\overrightarrow{v'}$ is not simply proportional to $\overrightarrow{p'}$. The time derivatives of the momentum components read:

$$\dot{p}'_x = -\frac{\partial H}{\partial x'} = -\frac{\partial U}{\partial x'} + \omega p'_y \tag{7.108}$$

$$\dot{p}'_y = -\frac{\partial H}{\partial y'} = -\frac{\partial U}{\partial y'} - \omega p'_x \tag{7.109}$$

$$\dot{p}'_z = -\frac{\partial H}{\partial z'} = -\frac{\partial U}{\partial z'}. \tag{7.110}$$

The combination of the equations above gives the well-known equations of motion:

$$\ddot{x}' - 2\omega\dot{y}' - \omega^2 x' = \frac{F_{x'}}{m} \tag{7.111}$$

$$\ddot{y}' + 2\omega\dot{x}' - \omega^2 y' = \frac{F_{y'}}{m} \tag{7.112}$$

$$\ddot{z}' = \frac{F_{z'}}{m} \tag{7.113}$$

in which by default **Coriolis** and **centripetal acceleration** appear. For the explicit proof one uses $\overrightarrow{\omega} = \omega\ \overrightarrow{e_z}$ and evaluates the Coriolis acceleration $2\overrightarrow{\omega} \times \overrightarrow{v'}$ as well as the centripetal acceleration $\overrightarrow{\omega} \times (\overrightarrow{\omega} \times \overrightarrow{r'})$ for each component.

Note: Examples 2 and 3 show that the **canonical** momentum, e.g. $p_x = \partial L/\partial v_x$, has to be distinguished from the **mechanical** momentum mv_x.

7.3 Symmetry and Conservation Laws

7.3.1 Cyclic Variables

If the Lagrangian function $L(q_i, \dot{q}_i, t)$ does not depend on the generalized coordinate q_C, i.e.

$$\boxed{\frac{\partial L}{\partial q_C} = 0, \qquad (7.114)}$$

the associated Lagrange equation gives

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_C}\right) = 0. \qquad (7.115)$$

The generalized momentum p_C is therefore a constant of motion,

$$\boxed{p_C = \frac{\partial L}{\partial \dot{q}_C} = \text{const.} \qquad (7.116)}$$

Generalized coordinates with the property (7.114) are denoted as **cyclic variables**.

Example: For a particle in a central field the lagrangian reads in spherical coordinates (r, ϑ, φ) :

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\vartheta}^2 + r^2 \sin^2\vartheta \; \dot{\varphi}^2) - U(r), \qquad (7.117)$$

and is independent of the angle φ, which is a cyclic variable. The associated generalized momentum thus is a conserved quantity,

$$p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = mr^2 \sin^2\vartheta \; \dot{\varphi} = l_z = \text{const.} \qquad (7.118)$$

7.3.2 Translation Invariance and Momentum Conservation

Due to the homogeneity of space the Lagrangian function of a closed system must be invariant with respect to translations, i.e.

$$L(\overrightarrow{r_i}, \overrightarrow{v_i}, t) = L(\overrightarrow{r_i} + \overrightarrow{a}, \overrightarrow{v_i}, t); \qquad (7.119)$$

where $\overrightarrow{a}$ is an arbitrary vector and the same for all particle displacement vectors. Since the translations form a continuous group it is sufficient to consider small shifts for which (by Taylor expansion) follows:

$$\sum_i \left(\frac{\partial L}{\partial x_i} a_x + \frac{\partial L}{\partial y_i} a_y + \frac{\partial L}{\partial z_i} a_z\right) = \sum_i \frac{\partial L}{\partial \overrightarrow{r_i}} \cdot \overrightarrow{a} = 0, \qquad (7.120)$$

i.e.

$$\sum_i \frac{\partial L}{\partial \overrightarrow{r_i}} = 0, \tag{7.121}$$

since $\overrightarrow{a}$ was arbitrary. From the Lagrange equations of motion we get:

$$\boxed{\frac{d}{dt}\left(\sum_i \frac{\partial L}{\partial \dot{\overrightarrow{r}}_i}\right) = \frac{d}{dt}\left(\sum_i \overrightarrow{p_i}\right) = 0,} \tag{7.122}$$

thus

$$\boxed{\overrightarrow{P} = \sum_{i=1}^{N} \overrightarrow{p_i} = \text{const},} \tag{7.123}$$

which corresponds to the **conservation of momentum**.

7.3.3 Rotational Invariance and Angular Momentum Conservation

Due to the isotropy of space we must have—in case of a closed system—for sufficiently small angles φ:

$$L(\overrightarrow{r_i}, \overrightarrow{v_i}, t) = L(\overrightarrow{r_i} + \varphi(\overrightarrow{u} \times \overrightarrow{r_i}), \overrightarrow{v_i} + \varphi(\overrightarrow{u} \times \overrightarrow{v_i}), t). \tag{7.124}$$

The vectors $\overrightarrow{r_i}$, $\overrightarrow{v_i}$ here are rotated by an angle φ around an arbitrary axis given by the unit vector $\overrightarrow{u}$. In analogy to the considerations in case of translation invariance, it follows by Taylor expansion:

$$\sum_i \frac{\partial L}{\partial \overrightarrow{r_i}} \cdot (\overrightarrow{u} \times \overrightarrow{r_i}) + \sum_i \frac{\partial L}{\partial \overrightarrow{v_i}} \cdot (\overrightarrow{u} \times \overrightarrow{v_i}) = 0, \tag{7.125}$$

or with the Langrange equations:

$$\sum_i \dot{\overrightarrow{p_i}} \cdot (\overrightarrow{u} \times \overrightarrow{r_i}) + \sum_i \overrightarrow{p_i} \cdot (\overrightarrow{u} \times \overrightarrow{v_i}) = 0. \tag{7.126}$$

With the cyclic invariance of the product, $\overrightarrow{a} \cdot (\overrightarrow{b} \times \overrightarrow{c}) = \overrightarrow{b} \cdot (\overrightarrow{c} \times \overrightarrow{a}) = \overrightarrow{c} \cdot (\overrightarrow{a} \times \overrightarrow{b})$, and the product rule equation (7.126) simplifies to

$$\frac{d}{dt}\left(\sum_i (\overrightarrow{r_i} \times \overrightarrow{p_i}) \cdot \overrightarrow{u}\right) = \sum_i \left((\overrightarrow{v_i} \times \overrightarrow{p_i}) \cdot \overrightarrow{u} + (\overrightarrow{r_i} \times \dot{\overrightarrow{p}}_i) \cdot \overrightarrow{u}\right) = 0. \tag{7.127}$$

Since the unit vector $\overrightarrow{u}$ can be chosen arbitrarily, we get

$$\boxed{\overrightarrow{L} = \sum_{i=1}^{N} \overrightarrow{l_i} = \sum_i (\overrightarrow{r_i} \times \overrightarrow{p_i}) = \text{const},} \tag{7.128}$$

i.e. **the angular momentum conservation**.

7.3.4 Time–Translation and Energy Conservation

The homogeneity of time allows us to set the time zero point arbitrarily. For a closed system the Lagrange function is invariant with respect to the transformation

$$t \to t + \tau \tag{7.129}$$

for any τ, i.e.

$$\frac{\partial L}{\partial t} = 0. \tag{7.130}$$

Using the Lagrange equations we get:

$$\frac{dL}{dt} = \sum_j \left(\frac{\partial L}{\partial q_j} \dot{q}_j + \frac{\partial L}{\partial \dot{q}_j} \ddot{q}_j \right) = \sum_j \left(\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_j} \right) \dot{q}_j + \frac{\partial L}{\partial \dot{q}_j} \ddot{q}_j \right) = \frac{d}{dt} \left(\sum_j \frac{\partial L}{\partial \dot{q}_j} \dot{q}_j \right); \tag{7.131}$$

accordingly

$$\frac{d}{dt} \left(L - \sum_j \frac{\partial L}{\partial \dot{q}_j} \dot{q}_j \right) = -\frac{d}{dt} H = 0 \tag{7.132}$$

and thus

$$\boxed{\sum_j \frac{\partial L}{\partial \dot{q}_j} \dot{q}_j - L = \sum_j p_j \dot{q}_j - L = H = \text{const.}} \tag{7.133}$$

The **Hamilton function** H of the system is therefore a **conserved quantity**. It is identical to the energy E of the system, if conservative forces and scleronome constraints exist. We then have:

$$L = T - U, \tag{7.134}$$

if U is the potential energy of the system, and

$$\boxed{\sum_j \frac{\partial L}{\partial \dot{q}_j} \dot{q}_j = 2T,} \tag{7.135}$$

such that

$$T - U - 2T = -H \tag{7.136}$$

or

$$\boxed{H = T + U = E.} \tag{7.137}$$

The special role of the Hamilton function H is also reflected in the canonical differential equations (7.83): the change of H with respect to a momentum p_i determines the time evolution of the associated coordinate q_i and vice versa.

For the **proof** of (7.135) we use the fact that for conservative forces the potential U does not depend on $\dot{q}_j$ such that

(7.138)

$$\frac{\partial L}{\partial \dot{q}_j} = \frac{\partial T}{\partial \dot{q}_j}.$$

For scleronome conditions we have

$$\vec{r}_i = \vec{r}_i(q_1, .., q_s) \tag{7.139}$$

and thus

$$\vec{v}_i = \sum_j \frac{\partial \vec{r}_i}{\partial q_j} \dot{q}_j, \tag{7.140}$$

where $\partial \vec{r}_i / \partial q_j$ is a function of the generalized coordinates q_l alone. The kinetic energy is therefore a quadratic form in the velocities $\dot{q}_j$:

$$T = \frac{1}{2} \sum_i m_i v_i^2 = \sum_{j,l} a_{jl} \; \dot{q}_j \dot{q}_l, \tag{7.141}$$

in which the coefficients a_{jl} only depend on the coordinates q_l. Then

$$\frac{\partial T}{\partial \dot{q}_r} = \sum_l a_{rl} \; \dot{q}_l + \sum_j a_{jr} \; \dot{q}_j = 2 \sum_l a_{rl} \; \dot{q}_l, \tag{7.142}$$

if one accounts for the symmetry of the coefficients $a_{jl} = a_{lj}$. With (7.142) the proof completes:

$$\sum_r \frac{\partial L}{\partial \dot{q}_r} \dot{q}_r = \sum_r \frac{\partial T}{\partial \dot{q}_r} \dot{q}_r = 2 \sum_{r,l} a_{rl} \; \dot{q}_r \dot{q}_l = 2\,T. \tag{7.143}$$

In summarizing this chapter we have defined generalized coordinates, that fulfill the constraints imposed on the system and also reduce the complexity of the problem by reducing the number of (linear independent) degrees of freedom. The equations of motion in generalized coordinates then have been derived from Newton's equations of motion. It is found that these equations can also be generated by a variational principle, which specifies a Lagrange function *L*, which is given by the difference between the kinetic and potential energy in case of conservative forces. Generalized momenta have been defined by the derivative of the Lagrange function with respect to the generalized velocities. Accordingly, if the Lagrange function does not depend on a specific coordinate, e.g. the azimuthal angle φ, the corresponding generalized momentum (here angular momentum) is a constant of motion. This suggested to transform the formulation to phase-space variables given by coordinates and their associated momenta, which was carried out by a Legendre transformation defining the Hamilton function *H*. In case of conservative forces the latter just gives the energy of the system in phase-space variables. The variational principle then can be reformulated in terms of Hamilton's (equivalent) variational principle, which gives the canonical equations of motion. The latter have been illustrated for a couple of examples. Furthermore, it was shown again that—for a closed system—the translational invariance leads to the conservation of the total momentum, the rotational invariance to the conservation of total angular momentum, and the invariance with respect to time translations to the conservation of the total energy.

W. Cassing, *Theoretical Physics Compact I*
https://doi.org/10.1007/978-3-031-95512-9_8

8. Applications of the Lagrange Formalism

Wolfgang Cassing[1] ✉
(1) University of Giessen, Gießen, Hessen, Germany

Applications of the Lagrange formalism will be given in this chapter for the motion of rigid bodies, which leads to the definition of an inertial tensor. The eigenvectors and eigenvalues of this tensor define the main axes of inertia and main moments of inertia, respectively. From the Lagrange function for the rigid body we will derive Euler's equation of motion, which will be studied for the case of a symmetric heavy gyroscope.

8.1 Motions of Rigid Bodies

As an explicit application of the Lagrange formalism we want to calculate the dynamics of a **rigid body**. A rigid body is a solid body whose mass elements form a solid and have a constant distance to each other such that they do not deform. Rigid bodies are also defined by the fact that only **translations** and **rotations** can be carried out. To describe rigid bodies we introduce two coordinate systems: an inertial system x_I, y_I, z_I (Fig. 8.1) and a body-fixed coordinate system x, y, z, which is firmly attached to the moving body (Fig. 8.2).

The motion of a rigid body consists of (i) a **translation**, in which the angular position of the body does not change and all mass points have the same velocity, and (ii) a **rotation** around a freely selectable coordinate origin O (Euler theorem). Since translations can be described by three coordinates and rotations by the axis of rotation and the size of the rotation angle, a freely moving rigid body has **six degrees of freedom**.

Since every motion of a rigid body consists of a translation and a rotation of the body-fixed coordinate system by O the velocity v_I of a point P—fixed to the body in the inertial system—is given by

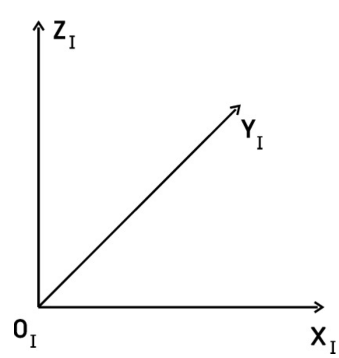

Fig. 8.1 Inertial system with axes x_I, y_I, z_I

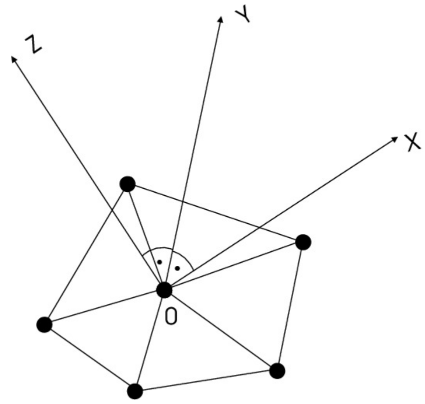

Fig. 8.2 Body-fixed coordinate system with axes x, y, z

$$\boxed{\vec{v}_I = \vec{v}_0 + \vec{\omega} \times \vec{r}} \tag{8.1}$$

where

- $\vec{v}_0 =$ the velocity of the coordinate origin O in the inertial system
- $\vec{\omega} =$ the angular velocity of the rigid body in the inertial system
- $\vec{r} = OP =$ the position vector of P in the rigid body coordinate system.

8.2 Kinetic Energy and Inertia Tensor

We assume that the rigid body consists of n mass points m_a. The kinetic energy then is:

$$T = \sum_{a=1}^{n} \frac{m_a}{2} v_{Ia}^2 = \sum_{a=1}^{n} \frac{m_a}{2} \left[\vec{v}_0 + (\vec{\omega} \times \vec{r}_a)\right]^2 \tag{8.2}$$

$$= \sum_{a=1}^{n} \frac{m_a}{2} \left(v_0^2 + \vec{v}_0 \cdot (\vec{\omega} \times \vec{r}_a) + ([\vec{\omega} \times \vec{r}_a)^2\right) \tag{8.3}$$

$$= \underbrace{\frac{M}{2} v_0^2}_{T_{\text{trans}}} + \underbrace{(\vec{v}_0 \times \vec{\omega}) \cdot \sum_{a=1}^{n} m_a \vec{r}_a}_{T_W} + \underbrace{\sum_{a=1}^{n} \frac{m_a}{2} (\vec{\omega} \times \vec{r}_a)^2}_{T_{\text{rot}}} \tag{8.4}$$

with the total mass

$$M := \sum_{a=1}^{n} m_a. \tag{8.5}$$

The first term is the translation energy T_{trans}, the third term the rotational energy T_{rot} and the middle term is an energy T_W, which is determined by translation and rotation both. If the rigid body is free, the coordinate origin O is best placed in the center of mass $\overrightarrow{S}$. Then $\sum_a m_a \overrightarrow{r_a} = 0$ and the energy T_W disappears, i.e.:

$$T = T_{\text{trans}} + T_{\text{rot}}. \tag{8.6}$$

The kinetic energy in this case is the sum of the kinetic translational energy of the mass M (located in the center of mass) and the rotational energy from the rotation around the center of mass.

If the rigid body is fixed in at least one point, the origin O of the body-fixed coordinate system is placed in one of these points and since $\overrightarrow{v_0} = 0$ we get:

$$T = T_{\text{rot}}. \tag{8.7}$$

The kinetic energy is equal to the rotational energy arising from the rotation around the fixed point.

We now recalculate the **rotation energy** within the body-fixed components ω_i and x_{ai}—with $i = 1, 2, 3$—of the vectors $\overrightarrow{\omega}$ and $\overrightarrow{r_a}$. With the notation

$$\overrightarrow{r_a} = (x_a, y_a, z_a) := (x_{a1}, x_{a2}, x_{a3}) \qquad a = 1, \ldots n. \tag{8.8}$$

and the identity

$$\left(\overrightarrow{a} \times \overrightarrow{b}\right)^2 = a^2 b^2 - \left(\overrightarrow{a} \cdot \overrightarrow{b}\right)^2 = \sum_{i,j=1}^{3} (a_i a_i b_j b_j - a_i b_i a_j b_j) \tag{8.9}$$

we obtain:

$$\begin{aligned} T_{\text{rot}} = \sum_{a=1}^{n} \frac{m_a}{2} \left(\overrightarrow{\omega} \times \overrightarrow{r_a}\right)^2 &= \sum_{a=1}^{n} \frac{m_a}{2} \sum_{i,j=1}^{3} [\omega_i \omega_i x_{aj} x_{aj} - \omega_i x_{ai} \omega_j x_{aj}] \\ &= \frac{1}{2} \sum_{a=1}^{n} m_a \sum_{i,j=1}^{3} \omega_i \omega_j \left[\sum_{k=1}^{3} x_{ak} x_{ak} \delta_{ij} - x_{ai} x_{aj}\right] \end{aligned} \tag{8.10}$$

with the **Kronecker symbol**

$$\delta_{ij} := \begin{cases} 1 & \text{for} \quad i = j \\ 0 & \text{for} \quad i \neq j \end{cases}. \tag{8.11}$$

In (8.10) the parameters (masses and positions) can now be derived from the projections of the angular velocity $\overrightarrow{\omega}$ on the body-fixed axes. To this aim we define the **inertia tensor** by

$$I_{ij} := \sum_{a=1}^{n} m_a \left[\sum_{k=1}^{3} x_{ak} x_{ak} \delta_{ij} - x_{ai} x_{aj}\right] \tag{8.12}$$

and obtain

$$T_{\rm rot} = \frac{1}{2}\sum_{i,j=1}^{3} I_{ij}\ \omega_i\omega_j. \tag{8.13}$$

It should be emphasized that the components ω_i are the **body-fixed** components of the angular velocity, i.e. the projections of $\vec{\omega}$ on the body fixed axes $x, y, z = x_1, x_2, x_3$.

Note:

- If the rigid body only rotates around one of its body-fixed axes, i.e. only a single component of $\vec{\omega}$ is different from zero, or if $I_{ij} = I\delta_{ij}$ holds, the above equation gives the more familiar result

$$T = \frac{1}{2} I\omega^2. \tag{8.14}$$

If the rigid body forms a continuous mass distribution we define the inertia tensor by

$$I_{ij} := \int \rho(x_1, x_2, x_3)\left[\sum_{k=1}^{3} x_k x_k \delta_{ij} - x_i x_j\right] dx_1 dx_2 dx_3 \tag{8.15}$$

with the mass density $\rho(x_1, x_2, x_3)$. To clarify, we present the inertia tensor I_{ij} (8.12) also in matrix notation:

$$I = \sum_{a=1}^{n} m_a \begin{pmatrix} y_a^2 + z_a^2 & -x_a y_a & -x_a z_a \\ -y_a x_a & x_a^2 + z_a^2 & -y_a z_a \\ -z_a x_a & -z_a y_a & x_a^2 + y_a^2 \end{pmatrix} \tag{8.16}$$

for n discrete masses m_a. In the case of a continuous mass density $\rho(x_1, x_2, x_3)$ we get with (8.15):

$$I = \int \rho(x_1, x_2, x_3) \begin{pmatrix} x_2^2 + x_3^2 & -x_1 x_2 & -x_1 x_3 \\ -x_2 x_1 & x_1^2 + x_3^2 & -x_2 x_3 \\ -x_3 x_1 & -x_3 x_2 & x_1^2 + x_2^2 \end{pmatrix} dx_1 dx_2 dx_3. \tag{8.17}$$

The diagonal elements of the inertia tensor are denoted by **moments of inertia**, the off-diagonal elements by **deviation moments**.

The inertia tensor by definition is symmetric (8.12):

$$I_{ij} = I_{ji} \tag{8.18}$$

and therefore, by introducing a new rotated coordinate system, can always be transformed to diagonal form. The corresponding axes will be denoted by **main axes of inertia** (eigenvectors), the diagonal elements $I_{ii} =: \lambda_i$ by **main moments of inertia** (eigenvalues).

The determination of the main axes of inertia, that pass through the center of mass, is simple for symmetrical bodies: one main axis of inertia coincides with the axis of symmetry; the other two main axes of inertia are orthogonal to it, but can be chosen arbitrarily. The rotational energy for the main axes of inertia is:

$$\boxed{T_{\rm rot} = \tfrac{1}{2}(\lambda_1\omega_1^2 + \lambda_2\omega_2^2 + \lambda_3\omega_3^2). \qquad (8.19)}$$

Def.: A rigid body is called

1. **rotator**, if it is one-dimensional and its mass points only lie on an axis, e.g. the *z* axis, such that $\lambda_1 = \lambda_2$; $\lambda_3 = 0$,
2. **asymmetric**, if all three main moments of inertia are different,
3. **symmetric**, if two main moments of inertia are equal,
4. **spherical top** if $\lambda_1 = \lambda_2 = \lambda_3$.

   Tops are not necessarily balls. For example, cubes are spherical tops and cylinders of height $h = \sqrt{3}r$, where *r* is the radius of the cylinder.

8.3 Angular Momentum

We assume again that the rigid body consists of *n* mass points m_a. The total angular momentum $\overrightarrow{L}_{tot}$ in the inertial frame then is

$$\overrightarrow{L}_{tot} = \textstyle\sum_{a=1}^{n} m_a(\overrightarrow{r}_{Ia} \times \overrightarrow{v}_i a). \qquad (8.20)$$

We denote the position vectors in the body-fixed coordinate system by $\overrightarrow{r}_a$, set

$$\overrightarrow{r}_{Ia} = \overrightarrow{r}_0 + \overrightarrow{r}_a, \qquad (8.21)$$

$$\overrightarrow{v}_{Ia} = \overrightarrow{v}_0 + \overrightarrow{\omega} \times \overrightarrow{r}_a \qquad (8.22)$$

and get:

$$\begin{aligned} \overrightarrow{L}_{\rm tot} &= M(\overrightarrow{r}_0 \times \overrightarrow{v}_0) + \overrightarrow{r}_0 \times \left[\overrightarrow{\omega} \times \left(\textstyle\sum_{a=1}^{n} m_a\overrightarrow{r}_a\right)\right] \\ &+ \left(\textstyle\sum_{a=1}^{n} m_a\overrightarrow{r}_a\right) \times \overrightarrow{v}_0 + \textstyle\sum_{a=1}^{n} m_a\left(\overrightarrow{r}_a \times (\overrightarrow{\omega} \times \overrightarrow{r}_a)\right). \end{aligned} \qquad (8.23)$$

Free system:

If the rigid body is not held fixed at any point, we place the coordinate origin O back to the center of mass $\vec{S} \Rightarrow \vec{r}_0 = \vec{r}_S$ and $\vec{v}_0 = \vec{v}_S$. Furthermore, from $\sum_a m_a \vec{r}_a = 0$ we obtain in the transformed system:

$$\vec{L}_{\text{tot}} = M\vec{r}_S \times \vec{v}_S + \underbrace{\sum_{a=1}^n m_a \vec{r}_a \times (\vec{\omega} \times \vec{r}_a)}_{\vec{L}} =: M(\vec{r}_S \times \vec{v}_S) + \vec{L}\ . \tag{8.24}$$

The total angular momentum $\vec{L}_{\text{tot}}$ is the sum of the term $M(\vec{r}_S \times \vec{v}_S)$, which is the **orbital angular momentum** of the center of mass motion with respect to the origin O_I, and the **intrinsic angular momentum** $\vec{L}$ for the rotation around the centroid $\vec{S}$.

Fixed system:

If the rigid body is fixed in at least one point, we place the coordinate origin O_I of the inertial system and the coordinate origin O of the body-fixed system in one of these fixed points and obtain due to $\vec{r}_0 = \vec{v}_0 = 0$:

$$\vec{L}_{\text{tot}} = \sum_{a=1}^n m_a \left(\vec{r}_a \times (\vec{\omega} \times \vec{r}_a)\right) =: \vec{L}. \tag{8.25}$$

For rotations about a fixed point the total angular momentum $\vec{L}_{\text{tot}}$ is equal to the intrinsic angular momentum $\vec{L}$, if both coordinate origins O_I and O are located in this point. With

$$\vec{r} \times (\vec{\omega} \times \vec{r}) = \vec{\omega}(\vec{r} \cdot \vec{r}) - \vec{r}(\vec{r} \cdot \vec{\omega}) \tag{8.26}$$

the **body-fixed** components of the intrinsic angular momentum $\vec{L}$ get the form

$$\boxed{L_i = \sum_{j=1}^3 \left(\sum_{a=1}^n m_a \left[\sum_{k=1}^3 x_{ak} x_{ak} \delta_{ij} - x_{ai} x_{aj}\right]\right)\omega_j = \sum_{j=1}^3 I_{ij}\ \omega_j \qquad i = 1, 2, 3.} \tag{8.27}$$

Using the main axes of inertia as body-fixed coordinate axes the **body-fixed** components of the intrinsic angular momentum $\vec{L}$ become:

$$\boxed{L_1 = \lambda_1 \omega_1, \qquad L_2 = \lambda_2 \omega_2, \qquad L_3 = \lambda_3 \omega_3.} \tag{8.28}$$

Accordingly, the intrinsic angular momentum $\vec{L}$ of a rigid body is generally not parallel to the angular velocity $\vec{\omega}$. Only when rotating around a main axis of inertia the angular momentum $\vec{L}$ and $\vec{\omega}$ have the same direction! The different directions of $\vec{L}$ and $\vec{\omega}$ are one of the reasons for the mathematical difficulty in the description of rigid bodies.

8.4 Euler's Equations

We now have to take a closer look at the angular momentum law. In general the inertia tensor is only constant in the body-fixed coordinate system, such that it is necessary for the equation of motion, i.e. primarily the time derivative of the angular momentum $\overrightarrow{L}_S$, to move back to the inertial system:

$$\dot{\overrightarrow{L}}_S = \frac{d}{dt}\left[\sum_{a=1}^{n} m_a \overrightarrow{r}_a \times (\overrightarrow{\omega} \times \overrightarrow{r}_a)\right] = \frac{d}{dt}\left[\sum_{i,j} I_{ij}\omega_j \overrightarrow{e}_i\right], \tag{8.29}$$

where $\omega_j = \overrightarrow{e}_j \cdot \overrightarrow{\omega}$ are the body-fixed coordinates of $\overrightarrow{\omega}$ and $\overrightarrow{e}_i$ are the basis vectors of the body-fixed coordinate system. With

$$\dot{\overrightarrow{e}}_i = \overrightarrow{\omega} \times \overrightarrow{e}_i \tag{8.30}$$

we get:

$$\boxed{\frac{d}{dt}\overrightarrow{L}_S = \dot{\overrightarrow{L}}_S = \sum_{i,j=1}^{3} I_{ij}\dot{\omega}_j \overrightarrow{e}_i + \overrightarrow{\omega} \times \sum_{i,j=1}^{3} I_{ij}\omega_j \overrightarrow{e}_i.} \tag{8.31}$$

The first term is the time derivative of the angular momentum for an observer in the body-fixed system and therefore the basis vectors $\overrightarrow{e}_i$ for him look constant. We denote this **body-fixed derivative** by $d_k\overrightarrow{L}_s/dt$ and get

$$\boxed{\frac{d}{dt}\overrightarrow{L}_S = \dot{\overrightarrow{L}}_S = \frac{d_k}{dt}\overrightarrow{L}_S + \overrightarrow{\omega} \times \overrightarrow{L}_S = \overrightarrow{N}_S,} \tag{8.32}$$

where the vectors $\overrightarrow{L}_S, \overrightarrow{\omega}, \overrightarrow{N}_S$ are expanded in the body-fixed basis and $\overrightarrow{N}_S$ denotes an external torque.

If the body-fixed axes are main axes of inertia, we find with $L_i = \lambda_i \omega_i$ by multiplying (8.31) or (8.32) by the basis vectors $\overrightarrow{e}_k$,

$$\overrightarrow{e}_k \cdot \frac{d}{dt}\overrightarrow{L}_S = \overrightarrow{e}_k \cdot \left(\sum_i \lambda_i \dot{\omega}_i \overrightarrow{e}_i\right) + \overrightarrow{e}_k \cdot \left(\overrightarrow{\omega} \times \sum_i \lambda_i \omega_i \overrightarrow{e}_i\right) = \overrightarrow{e}_k \cdot \overrightarrow{N}_S = N_k \tag{8.33}$$

for $k = 1, 2, 3$ the coupled nonlinear **Euler equations**

$$\boxed{\begin{aligned} \lambda_1\dot{\omega}_1 - (\lambda_2 - \lambda_3)\omega_2\omega_3 &= N_1 \\ \lambda_2\dot{\omega}_2 - (\lambda_3 - \lambda_1)\omega_3\omega_1 &= N_2 \\ \lambda_3\dot{\omega}_3 - (\lambda_1 - \lambda_2)\omega_1\omega_2 &= N_3. \end{aligned}} \tag{8.34}$$

Here ω_i and N_i are the projections of $\overrightarrow{\omega}$ and $\overrightarrow{N}$ on to the body-fixed coordinate axes $\overrightarrow{e_i}$, which must be the main axes of inertia.

As an **example** for the Euler equations (8.35) we examine the **force-free, symmetrical gyroscope**, i.e. $\overrightarrow{N_S} = 0$ and $\lambda_1 = \lambda_2$. The Eqs. (8.34) then simplify to

$$\lambda_1 \dot{\omega}_1 - (\lambda_1 - \lambda_3)\omega_2\omega_3 = 0$$

$$\lambda_1 \dot{\omega}_2 - (\lambda_3 - \lambda_1)\omega_3\omega_1 = 0$$

$$\lambda_3 \dot{\omega}_3 = 0. \tag{8.35}$$

From (8.34) it follows that ω_3 = const. and accordingly

$$\Omega = \frac{\lambda_3 - \lambda_1}{\lambda_1}\omega_3 = \text{const.} \tag{8.36}$$

As a result we get a linear coupled system in the variables ω_1, ω_2, i.e. with (8.36)

$$\dot{\omega}_1 + \Omega\omega_2 = 0, \qquad \dot{\omega}_2 - \Omega\omega_1 = 0. \tag{8.37}$$

We form another time derivative of the first equation, insert the 2nd equation and get

$$\ddot{\omega}_1 + \Omega^2\omega_1 = 0. \tag{8.38}$$

The solution of (8.38) is

$$\omega_1(t) = A \cos\,(\Omega t + \alpha) \tag{8.39}$$

with a phase α to be determined by the initial conditions. We obtain the solution for $\omega_2(t)$ by integration of the 2nd equation in (8.37) using (8.39):

$$\omega_2(t) = A \sin\,(\Omega t + \alpha), \tag{8.40}$$

such that with $\omega_1^2(t) + \omega_2^2(t) = A^2$ the magnitude of $\overrightarrow{\omega}$ is constant. The free symmetrical top rotates with the frequency Ω around the figure axis.

8.5 The Euler Angles

Euler's equations only determine the projections of the angular velocity $\overrightarrow{\omega}(t) \cdot \overrightarrow{e_i} = \omega_i(t)$. We now introduce Euler angles, which determine the angular position, i.e. the orientation of the body-fixed coordinate system (and thus of the body) in the inertial system, very clearly.

The transition from the inertial system Σ_I to the rotated rigid system Σ is carried out by three rotations, as shown in Fig. 8.3, in the following order:

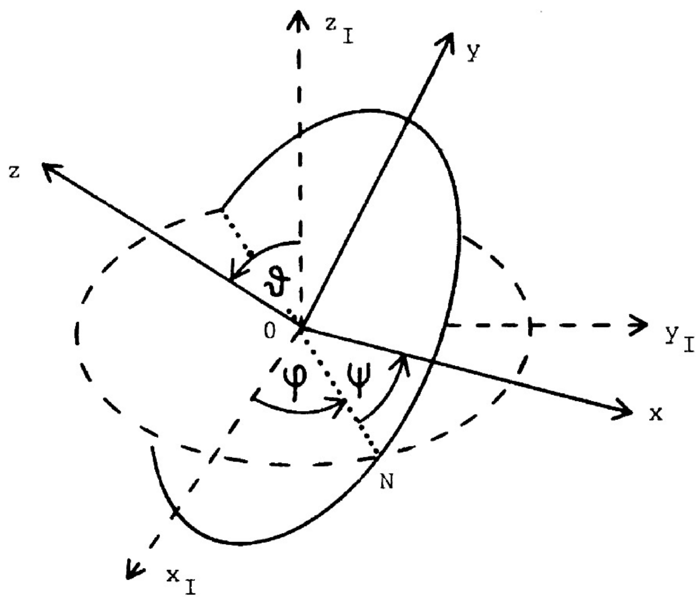

Fig. 8.3 Euler angles and rotations (see text)

1. **rotation by φ around the z_I-axis.** The x-axis goes over to the dotted 'nodal line' $0N$ and a new coordinate system $(\hat{x}, \hat{y}, \hat{z})$ emerges.
2. **rotation by ϑ around the nodal line $0N$.** The inertial z_I axis and the body-fixed z-axis then have the angle ϑ.
3. **rotation by ψ around the z-axis.** We get the body rigid coordinate system (x, y, z).

The Euler angles determine the orientation of the body-fixed coordinate system and thus also the rigid body relative to the inertial system: According to Fig. 8.3 the angles φ and ϑ give the position of the body-fixed z-axis in the inertial system. The angle ψ describes the rotation around the z-axis.

The angular velocity $\overrightarrow{\omega}$ is now written as the sum of the three Euler angular velocities $\overrightarrow{\omega}_\varphi, \overrightarrow{\omega}_\vartheta, \overrightarrow{\omega}_\psi$ as:

$$\overrightarrow{\omega} = \overrightarrow{\omega}_\varphi + \overrightarrow{\omega}_\vartheta + \overrightarrow{\omega}_\psi. \tag{8.41}$$

We project these three angular velocities onto the rigid body coordinate system in order to obtain the components $\omega_1, \omega_2, \omega_3$.

1. $\overrightarrow{\omega}_\varphi$ in the inertial system has the components

$$\overrightarrow{\omega}_{\varphi I} = \begin{pmatrix} 0 \\ 0 \\ \dot{\varphi} \end{pmatrix}, \tag{8.42}$$

and in the body-fixed system:

(8.43)

$$\vec{\omega}_\varphi = \dot{\varphi}\begin{pmatrix} \sin\psi \ \sin\vartheta \\ \cos\psi \ \sin\vartheta \\ \cos\vartheta \end{pmatrix}.$$

2. $\vec{\omega}_\vartheta$ in the coordinate system $(\hat{x}, \hat{y}, \hat{z})$ has the form

$$\vec{\hat{\omega}}_\vartheta = \begin{pmatrix} \dot{\vartheta} \\ 0 \\ 0 \end{pmatrix}, \tag{8.44}$$

such that in the body-fixed coordinate system we have:

$$\vec{\omega}_\vartheta = \dot{\vartheta}\begin{pmatrix} \cos\psi \\ -\sin\psi \\ 0 \end{pmatrix}. \tag{8.45}$$

3. For the angular velocity $\vec{\omega}_\psi$ we get:

$$\vec{\omega}_\psi = \begin{pmatrix} 0 \\ 0 \\ \dot{\psi} \end{pmatrix}. \tag{8.46}$$

The body-fixed components of $\vec{\omega}$ we obtain by adding the components:

$$\vec{\omega} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} = \begin{pmatrix} \dot{\varphi} \ \sin\vartheta\sin\psi \\ \dot{\varphi} \ \sin\vartheta\cos\psi \\ \dot{\varphi} \ \cos\vartheta \end{pmatrix} + \begin{pmatrix} \cos\psi \ \dot{\vartheta} \\ -\sin\psi \ \dot{\vartheta}\dot{\psi} \\ 0 \end{pmatrix}. \tag{8.47}$$

8.6 Lagrange Equations of the Rigid Body

With the preparatory work done the Lagrange function is set up easily. For a symmetric system with $\lambda_1 = \lambda_2$, whose body-fixed coordinate system coincides with the main axes of inertia, we get:

$$\boxed{T_{\text{rot}} = \frac{1}{2}\sum_{i,j} I_{ij}\omega_i\omega_j = \frac{1}{2}\sum_i \lambda_i\omega_i^2 = \frac{\lambda_1}{2}(\dot{\varphi}^2 \sin^2\vartheta + \dot{\vartheta}^2) + \frac{\lambda_3}{2}(\dot{\varphi}\cos\vartheta + \dot{\psi})^2. \quad (8.48)}$$

Example: **Heavy gyroscope**

A popular example for the application of the Lagrangian formalism is the symmetric top in the homogeneous gravity field, in which a point different from the center of mass on the axis of symmetry is fixed. Such a top is denoted by **heavy gyroscope**.

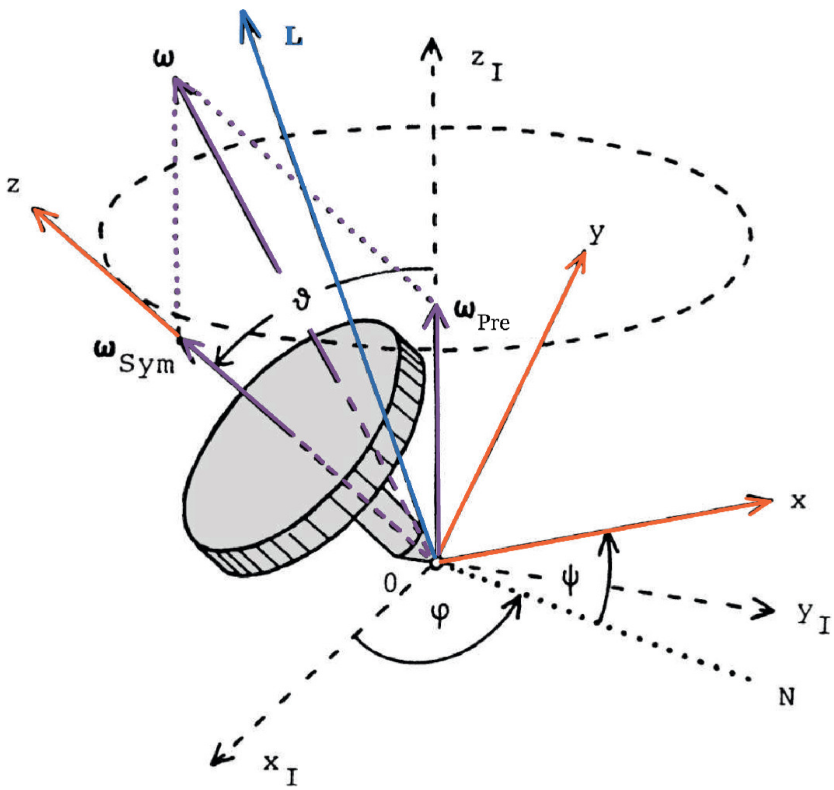

Fig. 8.4 Rotation of a heavy gyroscope (see text)

The zero points of the spatial and body-fixed coordinate systems are shown in Fig. 8.4 as well as the support point of the gyroscope. With the potential energy $U = mgl \cos \vartheta$ it's Lagrangian function according to (8.48) is:

$$L = T - U = \tfrac{\lambda_1}{2}(\dot{\varphi}^2 \sin^2 \vartheta + \dot{\vartheta}^2) + \tfrac{\lambda_3}{2}(\dot{\varphi} \cos \vartheta + \dot{\psi})^2 - mgl \cos \vartheta. \tag{8.49}$$

Here $\lambda_1 = \lambda_2, \lambda_3$ are the main moments of inertia for rotations around the support point, *m* is the mass of the top and *l* is the distance of the center of mass from the support point.

It follows that the angles φ, ψ, which represent the rotations about the z_I– and the *z*-axis, are cyclic and their momenta are conserved quantities:

$$p_\varphi = \tfrac{\partial L}{\partial \dot{\varphi}} = \lambda_1 \sin^2 \vartheta \dot{\varphi} + \lambda_3(\dot{\varphi} \cos \vartheta + \dot{\psi}) \cos \vartheta = \text{const.} \tag{8.50}$$

p_φ is the space-fixed z_I component of the angular momentum $\overrightarrow{L}$ and

$$p_\psi = \frac{\partial L}{\partial \dot{\psi}} = \lambda_3(\dot{\varphi} \cos \vartheta + \dot{\psi}) = \lambda_3 \omega_3 = \text{const.} \tag{8.51}$$

In the general case of a non-symmetric gyroscope with $\lambda_1 \neq \lambda_2 \neq \lambda_3$ we get the Lagrangian function L by substituting (8.47) into (8.48), which leads to somewhat more lengthy expressions, since L then consists of the angles ϑ, ψ and the time derivatives $\dot{\varphi}, \dot{\vartheta}, \dot{\psi}$ explicitly, only the variable φ is cyclic, if the potential does not depend on φ, i.e. $U \neq U(\varphi)$. The Lagrange equations of motion are modified correspondingly and their solutions become more subtle.

In summarizing this chapter we have given applications of the Lagrange formalism for the motion of rigid bodies, which lead to the definition of an inertial tensor. The eigenvectors and eigenvalues of this tensor define the main axes of inertia and main moments of inertia, respectively. From the Lagrange function for the rigid body we have derived Euler's equation of motion and solved these for the case of a symmetric heavy gyroscope.

W. Cassing, *Theoretical Physics Compact I*
https://doi.org/10.1007/978-3-031-95512-9_9

9. Dynamics in Phase Space

Wolfgang Cassing[1]
(1) University of Giessen, Gießen, Hessen, Germany

Although the Lagrange formalism is a convenient method to tackle complex problems, it is of advantage to formulate the dynamics in phase-space variables, i.e. in generalized coordinates and generalized momenta. In this case the time evolution of an observable, that not explicitly depends on time, is given by **Poisson brackets** which are determined by the derivative of the observable and the Hamiltonian with respect to the phase-space variables. The elementary Poisson bracket between generalized coordinates and generalized momenta will turn out to be unity for associated pairs and their time evolution is given by the Poisson bracket with the Hamilton function, i.e. by the canonical equations of motion. The Poisson brackets thus allow for an algebraic formulation of the dynamics. However, the choice of generalized coordinates is not unique and invertible transformations between the coordinates are allowed, too. But not all transformations are meaningful, since some transformations may lead to equations of motion that are no longer canonical. Allowed transformations then will be given by point transformations and extended canonical transformations, that keep the equations of motion **canonical invariant**. Furthermore, the **elementary Poisson brackets** will be shown to be invariant with respect to canonical transformations such that a formulation of classical mechanics is achieved, which is independent on the choice of the generalized coordinates. This will pave the way to quantum mechanics, where the Poisson brackets will be replaced by commutators of operators in an abstract Hilbert space. This also will lead to a rigid formulation of statistical mechanics, where the physical system—in equilibrium—is described by ensembles with properties, that are defined by expectation values of conserved quantities and their fluctuations.

9.1 Temporal Change of an Observable

We here attempt to 'directly' determine the temporal change of an observable

$$O = O(q_i, p_i; t) \tag{9.1}$$

of the system under consideration, such as the energy, the momentum, the magnetic moment in an external field etc. We do this by writing the total time-derivative of O as

$$\frac{dO}{dt} = \sum_{i=1}^{s}\left(\frac{\partial O}{\partial q_i}\dot{q}_i + \frac{\partial O}{\partial p_i}\dot{p}_i\right) + \frac{\partial O}{\partial t} \tag{9.2}$$

and use the canonical equations:

$$\frac{dO}{dt} = \sum_{i=1}^{s}\left(\frac{\partial O}{\partial q_i}\frac{\partial H}{dp_i} - \frac{\partial O}{\partial p_i}\frac{\partial H}{dq_i}\right) + \frac{\partial O}{\partial t}. \tag{9.3}$$

Equation (9.3) can also be written in terms of **Poisson brackets** defined by:

$$\{u, v\} = \sum_{i=1}^{s}\left(\frac{\partial u}{\partial q_i}\frac{\partial v}{dp_i} - \frac{\partial u}{\partial p_i}\frac{\partial v}{dq_i}\right). \tag{9.4}$$

This leads to the shorthand form:

$$\frac{dF}{dt} = \{F, H\} + \frac{\partial F}{\partial t}. \tag{9.5}$$

Special cases:

(i) $O = H$, then (9.5) reads

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = 0, \tag{9.6}$$

if H does not explicitly depend on time t, i.e. for a closed system H is constant.

(ii) **Canonical equations**

For $O = q_i$ we get:

$$\dot{q}_i = \{q_i, H\} = \frac{\partial H}{\partial p_i}, \tag{9.7}$$

since

$$\frac{\partial q_i}{\partial q_j} = \delta_{ij}; \qquad \frac{\partial q_i}{\partial p_j} = 0. \tag{9.8}$$

For $O = p_i$ we obtain:

$$\dot{p}_i = \{p_i, H\} = -\frac{\partial H}{\partial q_i}, \tag{9.9}$$

due to

$$\frac{\partial p_i}{\partial p_j} = \delta_{ij}; \qquad \frac{\partial p_i}{\partial q_j} = 0. \tag{9.10}$$

9.2 Properties of Poisson Brackets

The Poisson brackets defined in (9.4) are important not only in context with the time evolution of an observable, they also allow to formulate classical mechanics in a form where the connection with quantum mechanics can be clearly demonstrated. We therefore give a series of important rules below which simplify the calculation of Poisson brackets:

(i) **antisymmetry**

$$\{u, v\} = -\{v, u\} \tag{9.11}$$

(ii) **linearity**

$$\{u, v + w\} = \{u, v\} + \{u, w\} \tag{9.12}$$

(iii) **product rule**

$$\{u, vw\} = v\{u, w\} + \{u, v\}w \tag{9.13}$$

(iv) **Jacobi identity**

$$\{u, \{v, w\}\} + \{v, \{w, u\}\} + \{w, \{u, v\}\} = 0. \tag{9.14}$$

The proofs for (9.11)–(9.14) follow directly from the definition (9.4) and the standard rules of differentiation.

Examples:

(1.) **Canonical conjugate variable** q_i, p_i are distinguished because

$$\{q_i, q_j\} = 0; \qquad \{p_i, p_j\} = 0; \qquad \{q_i, p_j\} = \delta_{ij}. \tag{9.15}$$

(2.) **Angular momentum**: The following holds for the components of the angular momentum:

$$\boxed{\{L_1, L_2\} = L_3; \qquad \{L_3, L_1\} = L_2; \qquad \{L_2, L_3\} = L_1,} \tag{9.16}$$

as can be easily proven using (9.15). The quantum mechanical analogue of (9.16) is the basis for the quantization of angular momentum!

(3.) **Conserved quantities**: If an observable G is not explicitly dependent on time t, we have

$$\frac{dG}{dt} = \{G, H\} = 0, \tag{9.17}$$

and thus G= const., if $\{G, H\} = 0$. The importance of Poisson brackets is, that they provide an algebraic formulation for the dynamics of physical systems and allow for a formal 'introduction' to quantum mechanics, in which the conjugate variables (q_l, p_l) are replaced by **operators** in an abstract **Hilbert space** (see quantum mechanics).

(4.) **The harmonic oscillator**:

A completely algebraic solution is possible e.g. for the harmonic oscillator, i.e. for the Hamilton function

$$H(q,p) = \frac{p^2}{2m} + \frac{m}{2}\omega_0^2 q^2. \tag{9.18}$$

With (9.7) we get $\dot{q}$ using (9.15):

$$\dot{q} = \{q, H\} = \{q, \frac{p^2}{2m} + \frac{m}{2}\omega_0^2 q^2\} = \frac{1}{2m}\{q, p^2\}$$

$$= \frac{1}{2m}(p\{q,p\} + \{q,p\}p) = \frac{1}{2m}(p+p) = \frac{p}{m} \tag{9.19}$$

and $\dot{p}$ as:

$$\dot{p} = \{p, H\} = \{p, \frac{p^2}{2m} + \frac{m}{2}\omega_0^2 q^2\} = \frac{m\omega_0^2}{2}\{p, q^2\}$$

$$= \frac{m\omega_0^2}{2}(q\{p,q\} + \{p,q\}q) = \frac{m\omega_0^2}{2}(-q-q) = -m\omega_0^2 q. \tag{9.20}$$

Together:

$$\ddot{q} = \frac{\dot{p}}{m} = -\omega_0^2 q \quad \text{or} \quad \ddot{q} + \omega_0^2 q = 0, \tag{9.21}$$

i.e. an equation for a vibration with frequency ω_0.

9.3 Canonical Transformations

We now want to investigate the conditions, that the Lagrange equations and canonical equations of motion do not change for transformations of the $2s$ coordinates of a physical system, i.e. are **form invariant**.

9.3.1 Point Transformations

When formulating Lagrangian dynamics we have introduced generalized coordinates q_l such that constraints imposed by the system are met identically. However, the choice of generalized coordinates q_l for many-body systems is by no means clear and one can choose different coordinate systems. The question then arises, if the dynamics are invariant under **point transformations**

$$q_i \to Q_i(q_l;t), \qquad\qquad l = 1,..,s. \tag{9.22}$$

As an example for such a point transformation we mention again the transformation of cartesian coordinates to spherical coordinates:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \to \begin{pmatrix} r(x,y,z) \\ \vartheta(x,y,z) \\ \varphi(x,y,z) \end{pmatrix}. \tag{9.23}$$

On the other hand, we are interested in an 'optimal' set of coordinates Q_j in which **all cyclic variables of the system** occur explicitly.

We now show that the Lagrange equations are indeed **form-invariant** with respect to point transformations (9.22), i.e.

$$L(q_i,\dot{q}_i;t) \to L'(Q_i,\dot{Q}_i;t) = L(q_i(Q_j,t),\dot{q}_i(Q_j,\dot{Q}_j,t);t) \tag{9.24}$$

in the sense:

$$\boxed{\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \quad \Longleftrightarrow \frac{d}{dt}\frac{\partial L'}{\partial \dot{Q}_j} - \frac{\partial L'}{\partial Q_j} = 0.} \tag{9.25}$$

For the proof we calculate

$$\frac{\partial L'}{\partial Q_j} = \sum_{i=1}^s \frac{\partial L}{\partial q_i}\frac{\partial q_i}{\partial Q_j} = \sum_{i=1}^s a_{ij}\frac{\partial L}{\partial q_i} \tag{9.26}$$

with the $s \times s$ transformation matrix

(9.27)

$$a_{ij} = \frac{\partial q_i}{\partial Q_j}.$$

In anlogy we obtain for the momenta P_j with (9.27):

$$P_j = \frac{\partial L'}{\partial \dot{Q}_j} = \sum_{i=1}^{s} \frac{\partial L}{\partial \dot{q}_i} \frac{\partial \dot{q}_i}{\partial \dot{Q}_j} = \sum_{i=1}^{s} p_i \frac{\partial \dot{q}_i}{\partial \dot{Q}_j} = \sum_{i=1}^{s} p_i \frac{\partial q_i}{\partial Q_j} = \sum_{i=1}^{s} a_{ij}\, p_i, \tag{9.28}$$

i.e. with (9.26)

$$\frac{d}{dt} P_j = \sum_{i=1}^{s} a_{ij} \frac{d}{dt} p_i = \frac{\partial L'}{\partial Q_j} = \sum_{i=1}^{s} a_{ij} \frac{\partial L}{\partial q_i}. \tag{9.29}$$

The form invariance now follows from the fact, that the Lagrange equations in the coordinates q_l and Q_j emerge from each other by multiplying with an invertible $s \times s$ matrix $(a)_{ij}$ whose determinant is $\neq 0$.

For the Hamilton function $H'(Q_i, P_i; t)$ we get

$$\boxed{H(q_i, p_i; t) \to H'(Q_i, P_i; t) = \sum_i \dot{Q}_i P_i - L'(Q_i, \dot{Q}_i; t)} \tag{9.30}$$

and according to the variational principle (7.85) the equations of motion

$$\boxed{\dot{Q}_i = \frac{\partial H'}{\partial P_i}; \qquad \dot{P}_i = -\frac{\partial H'}{\partial Q_i}.} \tag{9.31}$$

Obviously the form of the equations of motion (9.31) is invariant with respect to a point transformation of the form (9.23).

9.3.2 Examples

Free particle in a plane

We restrict ourselves to the transformation in the (x, y) plane for a free particle of mass m, i.e. $\dot{z} = 0$:

$$q_i = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \to Q_i = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \\ z \end{pmatrix}. \tag{9.32}$$

The Lagrange function L then reads

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) \to L'(\dot{x}(r, \varphi, z, \dot{r}, \dot{\varphi}, \dot{z}), \dot{y}(r, \varphi, z, \dot{r}, \dot{\varphi}, \dot{z}); t) \tag{9.33}$$

and we obtain with

(9.34)

$$\dot{x} = \dot{Q}_1 = \tfrac{d}{dt}(r \cos \varphi) = \dot{r} \cos \varphi - r\dot{\varphi} \sin \varphi,$$

$$\dot{y} = \dot{Q}_2 = \tfrac{d}{dt}(r \sin \varphi) = \dot{r} \sin \varphi + r\dot{\varphi} \cos \varphi \tag{9.35}$$

the Lagrangian function

$$L' = \tfrac{m}{2}(\dot{r}^2 + r^2\dot{\varphi}^2) = L'(Q_i, \dot{Q}_i; t). \tag{9.36}$$

The momenta $P_i = \partial L'/\partial \dot{Q}_i$ result in

$$P_r = \tfrac{\partial L'}{\partial \dot{r}} = m\dot{r}, \qquad P_\varphi = \tfrac{\partial L'}{\partial \dot{\varphi}} = mr^2\dot{\varphi}. \tag{9.37}$$

The Hamilton function H' follows from (9.30)

$$H' = \dot{r}P_r + \dot{\varphi}P_\varphi - L'(\dot{r}, r, \dot{\varphi}) = \tfrac{P_r^2}{2m} + \tfrac{P_\varphi^2}{2mr^2}. \tag{9.38}$$

The equations of motion according to (9.31) are:

$$\dot{r} = \tfrac{\partial H'}{\partial P_r} = \tfrac{P_r}{m};\ \dot{\varphi} = \tfrac{\partial H'}{\partial P_\varphi} = \tfrac{P_\varphi}{mr^2};\ \dot{P}_r = -\tfrac{\partial H'}{\partial r} = \tfrac{P_\varphi^2}{mr^3};\ \dot{P}_\varphi = -\tfrac{\partial H'}{\partial \varphi} = 0, \tag{9.39}$$

i.e. the variable φ is cyclic.

Free particle in a rotating reference system
A particle of mass *m* continues to move in a system, which additionally rotates around the *z* axis with the angular velocity *ω*. We introduce the following new coordinates as:

$$r \to R = r; \qquad \varphi \to \Phi = \varphi + \omega t, \tag{9.40}$$

where the new coordinate Φ now explicitly depends on time *t*. The Lagrangian function $L''(R, \dot{R}, \Phi, \dot{\Phi}; t)$ then reads with

$$\dot{R} = \dot{r}; \qquad \dot{\Phi} = \dot{\varphi} + \omega \tag{9.41}$$

as:

$$L''(R, \dot{R}, \Phi, \dot{\Phi}; t) = \tfrac{m}{2}(\dot{R}^2 + R^2(\dot{\Phi} - \omega)^2). \tag{9.42}$$

With the momenta

$$P_r = \tfrac{\partial L''}{\partial \dot{R}} = m\dot{R}, \qquad P_\Phi = \tfrac{\partial L''}{\partial \dot{\Phi}} = mR^2(\dot{\Phi} - \omega) \tag{9.43}$$

the new Hamilton function H'' results in

$$H'' = \dot{R}P_R + \dot{\Phi}P_\Phi - L''(\dot{R}, R, \dot{\Phi}) = \frac{P_R^2}{2m} + \frac{P_\Phi^2}{2mR^2} + \omega P_\Phi. \tag{9.44}$$

Note: With (9.44) it becomes clear from the additional term ωP_Φ that from

$$L'(Q_i, \dot{Q}_i;t) = L(q_i(Q_i;t), \dot{q}_i(Q_i, \dot{Q}_i;t);t) \tag{9.45}$$

in general it **does not follow** that the Hamilton function H' can be calculated from H by inserting $q(Q_i, P_i;t), p(Q_i, P_i;t)$, i.e. for explicitly time-dependent transformations

$$H'(Q_i, P_i;t) \neq H(q_i(Q_i, P_i;t), p_i(Q_i, P_i;t);t). \tag{9.46}$$

The equations of motion for the free particle in the rotating reference system with the Hamilton function (9.44) read:

$$\begin{aligned} \dot{R} &= \frac{\partial H''}{\partial P_R} = \frac{P_R}{m};\ \dot{\Phi} = \frac{\partial H''}{\partial P_\Phi} = \frac{P_\Phi}{mR^2} + \omega; \\ \dot{P}_R &= -\frac{\partial H''}{\partial R} = \frac{P_\Phi^2}{mR^3};\ \dot{P}_\Phi = -\frac{\partial H''}{\partial \Phi} = 0, \end{aligned} \tag{9.47}$$

which implies that the variable Φ is cyclic in this case.

9.4 Extended Canonical Transformations

So far we have considered point transformations of the form (9.22), which are just transformations of the coordinates q_i. In the Hamilton function $H(q_i, p_i;t)$, however, the variables q_i and p_i are independent (equal right) variables, such that we have to investigate general transformations of the form

$$\begin{pmatrix} q_i \\ p_i \end{pmatrix} \to \begin{pmatrix} Q_i(q_j, p_j;t) \\ P_i(q_j, p_j;t) \end{pmatrix}. \tag{9.48}$$

Example: The extended transformation

$$\begin{pmatrix} q_i \\ p_i \end{pmatrix} \to \begin{pmatrix} Q_i \\ P_i \end{pmatrix} = \begin{pmatrix} -p_i \\ q_i \end{pmatrix}, \tag{9.49}$$

which exchanges coordinates and momenta, is canonical, since with $H(q_i, p_i;t)$ the Hamilton function $H'(Q_i, P_i;t)$ is given by

$$H'(Q_i, P_i;t) = H(P_i, -Q_i;t). \tag{9.50}$$

The canonical equations of motion follow

$$\frac{\partial H'(Q_j,P_j;t)}{\partial P_i} = \frac{\partial H(P_j,-Q_j;t)}{\partial P_i} = \frac{\partial H(q_j,p_j;t)}{\partial q_i} = -\dot{p}_i = \dot{Q}_i; \tag{9.51}$$

$$\frac{\partial H'(Q_j,P_j;t)}{\partial Q_i} = \frac{\partial H(P_j,-Q_j;t)}{\partial Q_i} = -\frac{\partial H(q_j,p_j;t)}{\partial p_i} = -\dot{q}_i = -\dot{P}_i, \tag{9.52}$$

thus the form invariance of the canonical equations of motion is shown with respect to the transformation (9.49).

The example clarifies that generalized coordinates and generalized momenta are 'exchangeable' and therefore have equal rights. Both degrees of freedom become 'abstract' coordinates in the Hamilton function, which can be represented by $2s$ independent degrees of freedom in the $2s$-dimensional phase space.

General transformations (9.48) are described by a transformation $T(q_j, p_j;t)$, which should be arbitrary but invertible, i.e. the inverse transformation $T^{-1}(P_i, Q_i;t)$ gives

$$\begin{pmatrix} Q_i \\ P_i \end{pmatrix} \to \begin{pmatrix} q_i(Q_j, P_j;t) \\ p_i(Q_j, P_j;t) \end{pmatrix}. \tag{9.53}$$

The problem with general invertible transformations T, however, is that the Lagrange equations are no longer 'from invariant'. Hamilton's equations then also are no longer 'form invariant', i.e. have the form (9.31). We therefore must look for 'restrictions' on the transformation T, which generally ensure 'form invariance'.

First we define suitable transformations as follows: we denote a transformation T **canonical in the lower sense**, if for all Hamilton functions $H(q_i, p_i;t)$ a function $H'(Q_i, P_i;t)$ in the new variables P_i, Q_i exists such that the equations of motion are 'form invariant'.

To provide suitable conditions for such transformations, let's go back to the variation principle (7.85), where the variations

$$\boxed{\begin{aligned} \delta S &= \delta \int_{t_1}^{t_2} \left(\sum_{i=1}^{s} \dot{q}_i p_i - H(q_i, p_i;t)\right)\, dt = 0 \\ &= \delta \int_{t_1}^{t_2} \left(\sum_{i=1}^{s} \dot{Q}_i P_i - H'(Q_i, P_i;t)\right)\, dt \end{aligned} \qquad (9.54)}$$

vanish for arbitrary interval boundaries t_1, t_2. We recall that the variational problem (9.54) leads directly to Hamilton's (canonical) equations of motion. The connection becomes immediately apparent, when we consider in addition to the actual trajectory $(q_i(t), p_i(t))$ any neighboring trajectory $(q_i(t) + \epsilon\eta_i(t), p_i(t) + \epsilon\kappa_i(t))$, where the functions η_i and κ_i must be linear independent since the q_i, p_i are also linear independent. The derivative of the action $S(\epsilon)$ with respect to ϵ leads (in the limit $\epsilon \to 0$) to:

$$\frac{dS}{d\epsilon} = \int_{t_1}^{t_2} \frac{d}{d\epsilon}\left(\sum_{i=1}^{s} [\dot{q}_i + \epsilon\dot{\eta}_i][p_i + \epsilon\kappa_i] - H(q_i + \epsilon\eta_i, p_i + \epsilon\kappa_i;t)\right)\, dt$$

$$= \int_{t_1}^{t_2} \left(\sum_{i=1}^{s} \dot{\eta}_i p_i + \dot{q}_i \kappa_i - \frac{\partial H}{\partial q_i} \eta_i - \frac{\partial H}{\partial p_i} \kappa_i \right) \, dt. \tag{9.55}$$

After partial integration of the term with $\dot{\eta}_i$ and consideration of the boundary conditions ($\eta_i(t_1) = \eta_i(t_2) = 0$) on the integration limits we get

$$\frac{dS}{d\epsilon} = \int_{t_1}^{t_2} \left(\sum_{i=1}^{s} [-\eta_i \dot{p}_i] + \dot{q}_i \kappa_i - \frac{\partial H}{\partial q_i} \eta_i - \frac{\partial H}{\partial p_i} \kappa_i \right) \, dt$$

$$= \int_{t_1}^{t_2} \sum_{i=1}^{s} \left[\left(-\dot{p}_i - \frac{\partial H}{\partial q_i} \right) \eta_i + \left(\dot{q}_i - \frac{\partial H}{\partial p_i} \right) \kappa_i \right] \, dt = 0. \tag{9.56}$$

Since the functions η_i, κ_i are arbitrary and linearly independent, the coefficients in the brackets (..) themselves must disappear, which just leads to the canonical equations of motion (9.31) in the variables (q_i, p_i).

We now come back to Equation (9.54). Since the variation is vanishing at the integration limits, the integrands differ–apart from an insignificant constant c—only by a total time differential of any continuous differentiable function F in the variables $q_i, p_i, Q_i, P_i; t$;

$$\begin{aligned} &\left(\sum_i \dot{q}_i p_i - H(q_i, p_i; t)\right) = \\ &c\left(\sum_i \dot{Q}_i P_i - H'(Q_i, P_i; t)\right) + \frac{d}{dt} F(q_i, p_i, Q_i, P_i; t), \end{aligned} \tag{9.57}$$

since the end points are kept during the variation, i.e.

$$\delta \int_{t_1}^{t_2} dt \frac{dF}{dt} = \delta(F(t_1) - F(t_2)) = 0. \tag{9.58}$$

After these preparatory remarks, we now **define** a transformation as **canonical**, if the constant c=1, i.e. if for any Hamilton function $H(q_i, p_i; t)$ a Hamilton function $H'(P_i, Q_i; t)$ exists with the property:

$$\boxed{\sum_{i=1}^{s} \left(\dot{q}_i p_i - P_i \dot{Q}_i \right) - H(q_i, p_i; t) + H'(Q_i, P_i; t) = \frac{d}{dt} F(q_i, p_i, Q_i, P_i; t). \quad (9.59)}$$

9.4.1 Generators of Canonical Transformations

The function introduced in (9.59) $F(q_i, p_i, Q_i, P_i; t)$ is an arbitrary (continuously differentiable) function of $4s + 1$ variables, where only $2s + 1$ are linear independent, since the number of degrees of freedom of the system is s and for each degree of freedom we need 2 independent variables; the time t is an additional parameter. Thus there are—except for linear combinations—only 6 different classes of **generating functions** with $2s + 1$ independent variables each:

$$F_1(q_i, Q_i; t), \;\; F_2(q_i, P_i; t), \;\; F_3(p_i, Q_i; t), \qquad (9.60)$$
$$F_4(p_i, P_i; t), F_5(q_i, p_i; t), \;\; F_6(Q_i, P_i; t).$$

Only F_5 is a function of the variables (q_i, p_i) alone such that (9.59) can be written in the form

$$\left(\sum_{i=1}^{s} \dot{q}_i p_i - H(q_i, p_i; t)\right) - \left(\sum_{i=1}^{s} P_i \dot{Q}_i - H'(Q_i, P_i; t)\right) = \frac{d}{dt} F_5(q_i, p_i; t)$$

$$= \sum_{i=1}^{s} \left(\dot{q}_i \frac{\partial F_5}{\partial q_i} + \dot{p}_i \frac{\partial F_5}{\partial p_i}\right) + \frac{\partial F_5}{\partial t}. \qquad (9.61)$$

The time derivative in the coordinate Q_i we can rewrite using the functional dependence on the $(q_i, p_i; t)$,

$$\dot{Q}_i = \sum_{k=1}^{s} \left(\frac{\partial Q_i}{\partial q_k} \dot{q}_k + \frac{\partial Q_i}{\partial p_k} \dot{p}_k\right) + \frac{\partial Q_i}{\partial t} \qquad (9.62)$$

and obtain from (9.61)

$$\sum_{i=1}^{s} \dot{q}_i p_i - \sum_{k=1}^{s} P_k \left(\sum_{i=1}^{s} \left[\frac{\partial Q_k}{\partial q_i} \dot{q}_i + \frac{\partial Q_k}{\partial p_i} \dot{p}_i\right] + \frac{\partial Q_k}{\partial t}\right) - H(q_i, p_i; t) + H'(Q_i, P_i; t)$$

$$= \sum_i \left(\dot{q}_i \frac{\partial F_5}{\partial q_i} + \dot{p}_i \frac{\partial F_5}{\partial p_i}\right) + \frac{\partial F_5}{\partial t}. \qquad (9.63)$$

Since the q_i, p_i are linearly independent, the quantities $\dot{q}_i, \dot{p}_i$ must also be linearly independent and thus the coefficients of the terms $\sim \dot{q}_i$ and $\sim \dot{p}_i$ vanish identically. By comparing the coefficients we get:

$$p_i - \sum_{k=1}^{s} P_k \frac{\partial Q_k}{\partial q_i} = \frac{\partial F_5}{\partial q_i}, \qquad (9.64)$$

$$-\sum_{k=1}^{s} P_k \frac{\partial Q_k}{\partial p_i} = \frac{\partial F_5}{\partial p_i}, \qquad (9.65)$$

$$H' = H + \sum_{k=1}^{s} P_k \frac{\partial Q_k}{\partial t} + \frac{\partial F_5}{\partial t}. \qquad (9.66)$$

The Eqs. (9.64) and (9.65) represent a system of coupled equations (of dimension $2s$), which can be solved for the $P_k(q_i, p_i; t), Q_k(q_i, p_i; t)$. The Hamilton function

$H'(Q_k, P_k;t)$ then follows from (9.66) by inserting the solutions $P_k(q_i, p_i;t), Q_k(q_i, p_i;t)$, where the partial time derivative of F_5 still can be chosen arbitrarily. **The generating function F_5 thus generates an infinite number of canonical transformations!** Without explicit proof we note that this holds true also for the generating function $F_6(Q_i, P_i;t)$, since it is also a function of the conjugate variables Q_i, P_i. The solution of the coupled system of equations (9.64) and (9.65), however, is quite complex since all equations include the functions P_k and Q_k in a nontrivial way.

Therefore, we will examine in the following the functions $F_1, \ldots, F_4$ and start with $F_1(q_i, Q_i;t)$. A transformation is called **canonical** if

$$\sum_{i=1}^{s} \dot{q}_i p_i - \sum_{i=1}^{s} P_i \dot{Q}_i - H(q_i, p_i;t) + H'(Q_i, P_i;t) = \frac{dF_1}{dt}$$

$$= \sum_{i=1}^{s} \left(\dot{q}_i \frac{\partial F_1}{\partial q_i} + \dot{Q}_i \frac{\partial F_1}{\partial Q_i} \right) + \frac{\partial F_1}{\partial t}. \tag{9.67}$$

Due to the linear independence of $\dot{q}_i$ and $\dot{Q}_i$ we obtain by a comparison of the coefficients

$$p_i = \frac{\partial F_1(q_i, Q_i;t)}{\partial q_i}, \tag{9.68}$$

$$P_i = -\frac{\partial F_1(q_i, Q_i;t)}{\partial Q_i}, \tag{9.69}$$

$$H' = H + \frac{\partial F_1(q_i, Q_i;t)}{\partial t}. \tag{9.70}$$

If the coordinates q_i, Q_i are linearly independent, the transformation to the coordinates p_i, P_i is **canonical**, if a function $F_1(q_i, Q_i;t)$ exists with the properties (9.68), (9.69) and (9.70).

As an **example** we calculate the transformation equations from the generating function

$$F_1(q, Q) = -\frac{Q}{q}. \tag{9.71}$$

According to (9.68) we get

$$p = \frac{\partial F_1(q, Q)}{\partial q} = \frac{Q}{q^2} \tag{9.72}$$

and with (9.69)

$$P = -\frac{\partial F_1(q,Q)}{\partial Q} = \frac{1}{q} = P(q,p). \tag{9.73}$$

With (9.72) this then results in

$$Q = pq^2 = Q(q,p), \tag{9.74}$$

i.e. the problem of the transformation equations from the variables (q, p) to the new variables (Q, P) is solved.

On the other hand, when knowing a transformation, e.g.

$$\begin{pmatrix} q \\ p \end{pmatrix} \rightarrow \begin{pmatrix} Q(q,p) \\ P(q,p) \end{pmatrix} = \begin{pmatrix} \ln p \\ -qp \end{pmatrix}, \tag{9.75}$$

we can calculate the generating function $F_1(q,Q)$. Equation (9.75) immediately gives

$$p = \exp\,(Q). \tag{9.76}$$

We start with (9.68) and integrate over dq, which gives F_1 in the form

$$F_1(q,Q;t) = \int p(q,Q)\ dq + g(Q;t) = q\ \exp\,(Q) + g(Q;t) \tag{9.77}$$

with any continuously differentiable function $g(Q;t)$. With (9.69) we get

$$P = -\frac{\partial F_1}{\partial Q} = -q\ \exp\,(Q) + \frac{\partial g(Q;t)}{\partial Q} = -qp(q,Q), \tag{9.78}$$

from which follows immediately:

$$\frac{\partial g(Q;t)}{\partial Q} = 0. \tag{9.79}$$

Thus the generating function $F_1 = q\ \exp\,(Q)$ is determined (except for an insignificant constant).

The **general procedure** for calculating the transformations $Q_j(q_i,p_i;t)$ and $P_j(q_i,p_i;t)$ is as follows: for a given $F_1(q_i,Q_i;t)$ one first calculates the s equations of motion for the p_i by differentiating the generators F_1 with respect to the q_i and solves the equations for the $Q_j(q_i,p_i;t)$. Then one calculates the derivatives of F_1 formally with respect to the Q_j and inserts the calculated $Q_j(q_i,p_i;t)$ into the expression obtained for the P_j, from which the transformations $P_j(q_i,p_i;t)$ finally result.

The generating function $F_2(q_i,P_i;t)$

We'll start with a function $\tilde{F}_2(q_i,P_i;t)$, which has the same **linear independent** variables as the function $F_2(q_i,P_i;t)$ (to be defined later). A transformation (9.48) then is canonical if:

$$\sum_{i=1}^{s} \dot{q}_i p_i - \sum_{i=1}^{s} P_i \dot{Q}_i - H(q_i, p_i;t) + H'(Q_i, P_i;t) = \frac{d\tilde{F}_2}{dt}$$

$$= \sum_{i=1}^{s} \left(\dot{q}_i \frac{\partial \tilde{F}_2}{\partial q_i} + \dot{P}_i \frac{\partial \tilde{F}_2}{\partial P_i} \right) + \frac{\partial \tilde{F}_2}{\partial t} =$$

$$\begin{aligned} &\sum_{i=1}^{s} \left[\dot{q}_i p_i - P_i \sum_{k=1}^{s} \left(\frac{\partial Q_i}{\partial q_k} \dot{q}_k + \frac{\partial Q_i}{\partial P_k} \dot{P}_k \right) - P_i \frac{\partial Q_i}{\partial t} \right] \\ &- H(q_i, p_i;t) + H'(Q_i, P_i;t) \end{aligned} \tag{9.80}$$

using the functional dependence $Q_i(q_k, P_k;t)$. Due to the linear independence of $\dot{q}_i$ and $\dot{P}_i$ we get (by a comparison of the coefficients)

$$p_i = \sum_{k=1}^{s} P_k \frac{\partial Q_k}{\partial q_i} + \frac{\partial \tilde{F}_2(q_j, P_j;t)}{\partial q_i} = \frac{\partial}{\partial q_i} \left[\tilde{F}_2 + \sum_{k=1}^{s} P_k Q_k \right], \tag{9.81}$$

$$0 = \sum_{k=1}^{s} P_k \frac{\partial Q_k}{\partial P_i} + \frac{\partial \tilde{F}_2(q_j, P_j;t)}{\partial P_i} = \frac{\partial}{\partial P_i} \left[\tilde{F}_2 + \sum_{k=1}^{s} P_k Q_k \right] - Q_i, \tag{9.82}$$

$$H' = H + \sum_{i=1}^{s} P_i \frac{\partial Q_i}{\partial t} + \frac{\partial \tilde{F}_2(q_j, P_j;t)}{\partial t} = H + \frac{\partial}{\partial t} \left[\tilde{F}_2 + \sum_{k=1}^{s} P_k Q_k \right], \tag{9.83}$$

where we also have used the linear independence of the variables (q_i, P_k), i.e. $\partial P_k / \partial q_i$=0.

The Eqs. (9.81), (9.82), (9.83) suggest to define a generating function $F_2(q_i, P_i;t)$ via

$$F_2(q_i, P_i;t) = \tilde{F}_2(q_i, P_i;t) + \sum_{k=1}^{s} P_k Q_k. \tag{9.84}$$

We then can write the Eqs. (9.81), (9.82), (9.83) in compact form:

$$p_i = \frac{\partial F_2(q_j, P_j;t)}{\partial q_i}, \tag{9.85}$$

$$Q_i = \frac{\partial F_2(q_j, P_j;t)}{\partial P_i}, \tag{9.86}$$

$$H' = H + \frac{\partial F_2(q_j, P_j;t)}{\partial t}. \tag{9.87}$$

Example: We calculate the generating function F_2 for the transformation

$$\begin{pmatrix} Q \\ P \end{pmatrix} = \begin{pmatrix} \ln p \\ -qp \end{pmatrix}. \tag{9.88}$$

With $p = -P/q$ we get by integration (9.85):

$$F_2(q, P) = \int p(P, q)dq + g(P) = -P \ln q + g(P) \tag{9.89}$$

with any continuous differentiable function $g(P)$. We now use (9.86) to get $g(P)$ via (9.89):

$$Q = \ln p = \frac{\partial F_2}{\partial P} = \frac{\partial[-P \ln q + g(P)]}{\partial P} = -\ln q + \frac{\partial g(P)}{\partial P}. \tag{9.90}$$

Integration of $\partial g(P)/\partial P$ over P yields (with $\ln(p) + \ln(q) = \ln(pq)$)

$$g(P) = \int \ln(pq)dP = \int \ln(-P)dP = P \ln(-P) - P. \tag{9.91}$$

Thus the generating function $F_2(q, P)$ reads

$$F_2(q, P) = -P \ln q + P \ln(-P) - P = P(\ln(-P/q) - 1)). \tag{9.92}$$

Relationship between the generators F_1 and F_2

From the defining equations for canonical transformations (9.67) and (9.80) we have immediately:

$$\frac{d}{dt}\left(F_1 - \tilde{F}_2\right) = 0 \qquad \text{or } F_1 = \tilde{F}_2 + \text{ const.}, \tag{9.93}$$

where the constant can be assumed to be 0 without any restrictions. With (9.84) we then get using (9.69):

$$\boxed{\begin{aligned} F_2(q_i, P_i;t) &= \tilde{F}_2(q_i, P_i;t) + \sum_{k=1}^{s} P_k Q_k \\ &= F_1(q_i, P_i;t) + \sum_{k=1}^{s}\left(-\frac{\partial F_1}{\partial Q_k}\right)Q_k \\ &= F_1(q_i, P_i;t) - \sum_{k=1}^{s} \frac{\partial F_1}{\partial Q_k} Q_k. \end{aligned}} \tag{9.94}$$

It thus turns out that the generating function F_2 is the **Legendre transform** of F_1.

9.4.2 Overview of the Generating Functions

In analogy to the previous considerations one finds that the generating functions F_3 and F_4 are also Legendre transforms of F_1:

$$F_3(p_i, Q_i;t) = F_1(q_i, Q_i;t) - \sum_{k=1}^{s} \frac{\partial F_1}{\partial q_k} q_k, \tag{9.95}$$

while $F_4(p_i, P_i;t)$ results from a double Legendre transformation:

$$\begin{aligned} F_4(p_i, P_i;t) &= F_1(q_i, Q_i;t) - \sum_{k=1}^{s} \left(\frac{\partial F_1}{\partial q_k} q_k + \frac{\partial F_1}{\partial Q_k} Q_k \right) \\ &= F_1(q_i, Q_i;t) + \sum_{k=1}^{s} (P_k Q_k - p_k q_k). \end{aligned} \tag{9.96}$$

The connections, that follow from the requirements (9.59) by a comparison of the coefficient, are given in Table 9.1 for the generators $F_1, \ldots, F_4$:

Table 9.1 Overview of the generating functions F_1, F_2, F_3 and F_4 and the definition of their variables

Overview			
$F_1(q, Q;t)$	$p = +\partial F_1/\partial q$	$P = -\partial F_1/\partial Q$	$H' = H + \partial F_1/\partial t$
$F_2(q, P;t)$	$p = +\partial F_2/\partial q$	$Q = +\partial F_2/\partial P$	$H' = H + \partial F_2/\partial t$
$F_3(p, Q;t)$	$q = -\partial F_3/\partial p$	$P = -\partial F_3/\partial Q$	$H' = H + \partial F_3/\partial t$
$F_4(p, P;t)$	$q = -\partial F_4/\partial p$	$Q = +\partial F_4/\partial P$	$H' = H + \partial F_4/\partial t$

Remark 1: From Table 9.1 it follows immediately that for time-independent transformations the Hamilton function itself is a **canonical invariant**, i.e. $H' = H$.

Remark 2: All point transformations $q_i \to Q_i(q_j;t)$ are canonical since there is a generating function

$$F_2(q_i, P_i;t) = \sum_{i=1}^{s} Q_i(q_j;t) P_i \tag{9.97}$$

with

$$p_i = \frac{\partial F_2}{\partial q_i} = \sum_{k=1}^{s} \frac{\partial Q_k}{\partial q_i}(q_j;t) P_k \tag{9.98}$$

and

$$Q_i = \frac{\partial F_2}{\partial P_i} = \sum_{k=1}^{s} \frac{\partial P_k}{\partial P_i}(q_j;t) Q_k. \tag{9.99}$$

As an **example** we consider the harmonic oscillator again,

$$H(q,p) = \frac{p^2}{2m} + \frac{m}{2}\omega_0^2 q^2, \tag{9.100}$$

and examine the canonical transformation generated by the function

$$F_1(q,Q) = \frac{m}{2}\omega_0 q^2 \cot(Q). \tag{9.101}$$

According to Table 9.1 we get:

$$p = \frac{\partial F_1}{\partial q} = m\omega_0 q \cot(Q); \qquad P = -\frac{\partial F_1}{\partial Q} = \frac{m\omega_0 q^2}{2\sin^2(Q)} \text{ or } q^2 = \frac{2P\sin^2(Q)}{m\omega_0}. \tag{9.102}$$

By simple transformations we obtain:

$$p = m\omega_0 q\frac{\cos(Q)}{\sin(Q)} = m\omega_0\frac{\cos(Q)}{\sin(Q)}\sqrt{\frac{2P}{m\omega_0}} \sin(Q) = \sqrt{2Pm\omega_0} \cos(Q) = p(P,Q). \tag{9.103}$$

$$q = \frac{p\sin(Q)}{m\omega_0\cos(Q)} = \sqrt{2Pm\omega_0} \cos(Q)\frac{\sin(Q)}{m\omega_0\cos(Q)} = \sqrt{\frac{2P}{m\omega_0}} \sin(Q) = q(P,Q). \tag{9.104}$$

The Hamilton function $H'(Q,P)$ in the new coordinates is given by:

$$H'(Q,P) = H(q(P,Q),p(Q,P)) = \frac{p^2}{2m} + \frac{m}{2}\omega_0^2 q^2 =$$

$$= \frac{2Pm\omega_0\cos^2(Q)}{2m} + \frac{m}{2}\omega_0^2\frac{2P}{m\omega_0} \sin^2(Q) = P\omega_0 \cos^2(Q) + P\omega_0 \sin^2(Q) = P\omega_0, \tag{9.105}$$

and the equations of motion in the coordinates P, Q are:

$$\dot{P} = -\frac{\partial H'}{\partial Q} = 0 \qquad \dot{Q} = \frac{\partial H'}{\partial P} = \omega_0. \tag{9.106}$$

These equations of motion show immediately that $P = H'/\omega_0 = P_0$ is a constant of motion, which is proportional to the energy $E = H'$. On the other hand, the solution for the angle variable follows immediately from the second equation,

$$Q(t) = \omega_0 t + \alpha, \tag{9.107}$$

where α denotes an arbitrary phase that has to specified by initial conditions. The solution is complete, when inserting the results for P and $Q(t)$ into the transformation formulae (9.103) and (9.104):

$$q(t) = \sqrt{\frac{2P_0}{m\omega_0}} \sin(\omega_0 t + \alpha), \tag{9.108}$$

$$p(t) = \sqrt{2P_0 m\omega_0} \cos{(\omega_0 t + \alpha)}. \tag{9.109}$$

The induced transformation by the generating function F_1 (9.101) thus allows for a simple solution of the oscillator problem.

9.4.3 Canonical Invariants

We denote quantities, which do not change with respect to canonical transformations, by **canonical invariants**. So far we have pointed out the invariance of the Hamilton function *H* as an example for time-independent canonical transformations, and the form invariance of Hamilton's equations of motion. We now will show that the formulation of the dynamics can be formulated **canonically invariant** with the help of the Poisson brackets (9.4) **for time-independent transformations**. We start with the.

Invariance of fundamental Poisson brackets

Let (q_i, p_i) and (Q_j, P_j) be two canonically conjugated sets of variables, for which both the Hamiltonian equations of motion hold with

$$\boxed{H'(Q_j, P_j) = H(q_i(Q_j, P_j), p_i(Q_j, P_j)). \quad (9.110)}$$

Then the following Poisson brackets are canonical invariants:

$$\boxed{\{Q_i, Q_j\}_{p,q} = 0;\ \{P_i, P_j\}_{p,q} = 0; \{Q_i, P_j\}_{p,q} = \delta_{ij}. \quad (9.111)}$$

To prove (9.111) we calculate the time derivative of Q_i,

$$\dot{Q}_i = \sum_{k=1}^{s} \left(\frac{\partial Q_i}{\partial q_k} \dot{q}_k + \frac{\partial Q_i}{\partial p_k} \dot{p}_k \right) = \sum_{k=1}^{s} \left(\frac{\partial Q_i}{\partial q_k} \frac{\partial H}{\partial p_k} - \frac{\partial Q_i}{\partial p_k} \frac{\partial H}{\partial q_k} \right) =$$

$$= \sum_{k,l=1}^{s} \left(\frac{\partial Q_i}{\partial q_k} \left[\frac{\partial H'}{\partial Q_l} \frac{\partial Q_l}{\partial p_k} + \frac{\partial H'}{\partial P_l} \frac{\partial P_l}{\partial p_k} \right] - \frac{\partial Q_i}{\partial p_k} \left[\frac{\partial H'}{\partial Q_l} \frac{\partial Q_l}{\partial q_k} + \frac{\partial H'}{\partial P_l} \frac{\partial P_l}{\partial q_k} \right] \right) =$$

$$= \sum_{k,l=1}^{s} \left(\frac{\partial H'}{\partial Q_l} \left[\frac{\partial Q_i}{\partial q_k} \frac{\partial Q_l}{\partial p_k} - \frac{\partial Q_i}{\partial p_k} \frac{\partial Q_l}{\partial q_k} \right] + \frac{\partial H'}{\partial P_l} \left[\frac{\partial Q_i}{\partial q_k} \frac{\partial P_l}{\partial p_k} - \frac{\partial Q_i}{\partial p_k} \frac{\partial P_l}{\partial q_k} \right] \right) =$$

$$= \textstyle\sum_{l=1}^{s} \left(-\dot{P}_l \{Q_i, Q_l\}_{p,q} + \dot{Q}_l \{Q_i, P_l\}_{p,q} \right) = \dot{Q}_i. \tag{9.112}$$

Consequently we must have:

$$\{Q_i, Q_l\}_{p,q} = 0; \qquad \{Q_i, P_l\}_{p,q} = \delta_{il}. \tag{9.113}$$

The still missing proof for $\{P_i, P_l\}_{p,q} = 0$ follows in analogy from the calculation of $\dot{P}_i$.

General Poisson brackets
We now want to show that the value of a Poisson bracket is independent of the set of canonical variables that are used as a basis. To this aim we consider any two phase-space functions *F* and *G* and two sets of canonical variables (q_i, p_i) and $(Q_j. P_j)$

$$\begin{pmatrix} q_l \\ p_l \end{pmatrix} = \begin{pmatrix} q_l(Q_j, P_j) \\ p_l(Q_j, P_j) \end{pmatrix}, \qquad \begin{pmatrix} Q_l \\ P_l \end{pmatrix} = \begin{pmatrix} Q_l(q_j, p_j) \\ P_l(q_j, p_j) \end{pmatrix}. \tag{9.114}$$

The Poisson bracket of *F* and *G* in the variables *q*, *p* then gives:

$$\{F, G\}_{p,q} = \sum_{j=1}^{s} \left(\frac{\partial F}{\partial q_j} \frac{\partial G}{\partial p_j} - \frac{\partial F}{\partial p_j} \frac{\partial G}{\partial q_j} \right) =$$

$$= \sum_{j,l=1}^{s} \left(\frac{\partial F}{\partial q_j} \left[\frac{\partial G}{\partial Q_l} \frac{\partial Q_l}{\partial p_j} + \frac{\partial G}{\partial P_l} \frac{\partial P_l}{\partial p_j} \right] - \frac{\partial F}{\partial p_j} \left[\frac{\partial G}{\partial Q_l} \frac{\partial Q_l}{\partial q_j} + \frac{\partial G}{\partial P_l} \frac{\partial P_l}{\partial q_j} \right] \right) =$$

$$= \sum_{l=1}^{s} \left(\tfrac{\partial G}{\partial Q_l} \{F, Q_l\}_{p,q} + \tfrac{\partial G}{\partial P_l} \{F, P_l\}_{p,q} \right). \tag{9.115}$$

Two intermediate results following immediately from (9.115) are:
(i) For $F = Q_k$ we get using the fundamental Poisson brackets:

$$\boxed{\{G, Q_k\}_{q,p} = -\tfrac{\partial G}{\partial P_k}.} \tag{9.116}$$

(ii) For $F = P_k$ the result is:

$$\boxed{\{G, P_k\}_{q,p} = \tfrac{\partial G}{\partial Q_k}.} \tag{9.117}$$

Inserting (9.116) and (9.117) into (9.115) we obtain the invariance of the Poisson bracket with respect to canonical transformations, since *F* and *G* have been chosen arbitrarily:

$$\boxed{\{F, G\}_{q,p} = \sum_{l=1}^{s} \left(\tfrac{\partial G}{\partial Q_l} \left[-\tfrac{\partial F}{\partial P_l} \right] + \tfrac{\partial G}{\partial P_l} \left[\tfrac{\partial F}{\partial Q_l} \right] \right) = \{F, G\}_{P,Q}.} \tag{9.118}$$

Accordingly we can omit the indices at the Poisson brackets, which have been introduced to specify the basic variables.

9.4.4 Criteria for Canonical Transformations

In practice the question often arises, if a specific transformation is canonical or not. This question often is not easy to answer if the associated explicit generating function is not known. For practical purposes, on the other hand, the following theorem is of great help:

An extended transformation (9.48) is **canonical**, if and only if the fundamental Poisson brackets are fulfilled in the new variables, i.e.

$$\{Q_i, Q_j\} = 0 = \{P_i, P_j\}; \qquad \{Q_i, P_j\} = \delta_{ij}. \tag{9.119}$$

We provide the proof for non-explicitly time-dependent transformations, i.e. for vanishing explicit time derivative of the generators $\partial F_k / \partial t = 0$, such that again:

$$H(q_i, p_i) = H'(Q_j, P_j) = H(q_i(Q_j, P_j), p_i(Q_j, P_j)). \tag{9.120}$$

Since according to Sect. 9.4.3 the Poisson brackets are invariant with respect to canonical transformations we choose, for the sake of simplicity, the variables q_i, p_i. The time derivative of Q_j and P_j then reads:

$$\dot{Q}_j = \{Q_j, H\}_{q,p} = \sum_{l=1}^{s} \left(\frac{\partial Q_j}{\partial q_l} \frac{\partial H}{\partial p_l} - \frac{\partial Q_j}{\partial p_l} \frac{\partial H}{\partial q_l} \right), \tag{9.121}$$

$$\dot{P}_j = \{P_j, H\}_{q,p} = \sum_{l=1}^{s} \left(\frac{\partial P_j}{\partial q_l} \frac{\partial H}{\partial p_l} - \frac{\partial P_j}{\partial p_l} \frac{\partial H}{\partial q_l} \right). \tag{9.122}$$

The partial derivatives of the Hamilton function can be rewritten as follows:

$$\frac{\partial H}{\partial p_l} = \sum_{k=1}^{s} \left(\frac{\partial H'}{\partial Q_k} \frac{\partial Q_k}{\partial p_l} + \frac{\partial H'}{\partial P_k} \frac{\partial P_k}{\partial p_l} \right). \tag{9.123}$$

$$\frac{\partial H}{\partial q_l} = \sum_{k=1}^{s} \left(\frac{\partial H'}{\partial Q_k} \frac{\partial Q_k}{\partial q_l} + \frac{\partial H'}{\partial P_k} \frac{\partial P_k}{\partial q_l} \right). \tag{9.124}$$

We insert (9.123) and (9.124) into Eq. (9.121),

$$\begin{aligned} \dot{Q}_j &= \{Q_j, H\}_{q,p} \\ &= \sum_{l,k=1}^{s} \left(\frac{\partial Q_j}{\partial q_l} \left(\frac{\partial H'}{\partial Q_k} \frac{\partial Q_k}{\partial p_l} + \frac{\partial H'}{\partial P_k} \frac{\partial P_k}{\partial p_l} \right) - \frac{\partial Q_j}{\partial p_l} \left(\frac{\partial H'}{\partial Q_k} \frac{\partial Q_k}{\partial q_l} + \frac{\partial H'}{\partial P_k} \frac{\partial P_k}{\partial q_l} \right) \right), \end{aligned} \tag{9.125}$$

and summarize as:

$$\dot{Q}_j = \{Q_j, H\}_{q,p} = \sum_{k=1}^{s}\left(\frac{\partial H'}{\partial Q_k}\{Q_j, Q_k\}_{q,p} + \frac{\partial H'}{\partial P_k}\{Q_j, P_k\}_{q,p}\right). \tag{9.126}$$

In the same way we find with (9.122):

$$\dot{P}_j = \{P_j, H\}_{q,p} = \sum_{k=1}^{s}\left(-\frac{\partial H'}{\partial Q_k}\{Q_k, P_j\}_{q,p} + \frac{\partial H'}{\partial P_k}\{P_j, P_k\}_{q,p}\right). \tag{9.127}$$

Hamilton's equations of motion

$$\dot{Q}_j = \frac{\partial H'}{\partial P_j}, \qquad \dot{P}_j = -\frac{\partial H'}{\partial Q_j} \tag{9.128}$$

thus hold, if and only if the fundamental Poisson brackets (9.119) are fulfilled in the new variables (q.e.d.).

The formulation of Newtonian dynamics in form of Poisson brackets, which are invariant with respect to canonical transformations and introduce **conjugate variables** by the fundamental Poisson brackets (9.119), allows for a simple transition to **quantum mechanics**.

9.5 Liouville's Theorem

Liouville's **theorem** provides an elegant introduction to **statistical mechanics**. In order to specify the state of a system of particles as a point in phase space exactly, one has to define (or measure) initial conditions for solving the canonical equations; for systems with a lot of particles ($N \sim 10^{23}$) this is practically impossible. Then a less precise (but for many questions still sufficient) description is to specify the **probability** $\rho(q_i, p_i;t)$ for a system to be at point (q_i, p_i) at time t in phase space. Knowing the probability $\rho(q_i, p_i;t)$ one can calculate the **expectation value** of an observable G as an average:

$$< G >= \int \rho(q_i, p_i;t)\ \ G(q_i, p_i;t) \prod_i dq_i dp_i \tag{9.129}$$

with the normalization

$$\int \rho(q_i, p_i;t) \prod_i dq_i dp_i = 1. \tag{9.130}$$

If the mean-square deviations $\Delta G^2 =< G^2 > - < G >^2$ are sufficiently small (which is the case for large numbers of particles) one can identify the average (9.129) with the macroscopic measurement.

The concept of the **ensemble** serves as an illustration of ρ in statistical mechanics: One replaces the actual system, whose initial conditions are imprecise (or incompletely known), by a sequence of many similar systems (**ensembles**) with different, but precisely specified initial conditions each, in accordance with the macroscopic knowledge about the actual system. Each member of the ensemble is represented by a point in phase space and the ensemble by a 'swarm' of points in phase space; their distribution is determined by the probability $\rho(q_i, p_i; t)$.

Following this idea we obtain the **Liouville equation** for the distribution function ρ , which reads:

$$\frac{d\rho}{dt} = \{\rho, H\} + \frac{\partial\rho}{\partial t} = 0. \tag{9.131}$$

To explain (9.131) we use the canonical equations and get

$$\{\rho, H\} = \sum_i \left(\frac{\partial\rho}{\partial q_i} \dot{q}_i + \frac{\partial\rho}{\partial p_i} \dot{p}_i \right) \tag{9.132}$$

according to the definition (9.4). Since now

$$\rho \sum_i \left(\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right) = \rho \sum_i \left(\frac{\partial^2 H}{\partial q_i \partial p_i} - \frac{\partial^2 H}{\partial p_i \partial q_i} \right) = 0, \tag{9.133}$$

we get

$$\{\rho, H\} = \sum_i \left(\frac{\partial}{\partial q_i} (\rho \dot{q}_i) + \frac{\partial}{\partial p_i} (\rho \dot{p}_i) \right), \tag{9.134}$$

thus

$$\sum_i \left(\frac{\partial}{\partial q_i} (\rho \dot{q}_i) + \frac{\partial}{\partial p_i} (\rho \dot{p}_i) \right) + \frac{\partial\rho}{\partial t} = 0 \tag{9.135}$$

due to (9.131).

Equation (9.135) can now be written as a **continuity equation in phase space**,

$$\frac{\partial\rho}{\partial t} + div(\rho\overrightarrow{v}) = 0 \tag{9.136}$$

with

$$\overrightarrow{v} = \begin{pmatrix} \dot{q}_i \\ \dot{p}_i \end{pmatrix} \tag{9.137}$$

as **velocity** in phase space and

$$div = \left(\frac{\partial}{\partial q_i}, \frac{\partial}{\partial pi} \right). \tag{9.138}$$

The **Liouville Theorem**—expressed in (9.131), (9.135) or (9.136)—then can be interpreted as the conservation of the number of the (ensemble representing) points in phase space (in analogy to the conservation of charge in electrodynamics): according to (9.136) the number of points in a certain area V_{Ph} of the phase space can only change if points of the 'swarm' move in or out.

Of particular interest for equilibrium thermodynamics is the case of a **stationary distribution**,

$$\frac{\partial \rho}{\partial t} = 0, \tag{9.139}$$

for which holds

$$\{\rho, H\} = 0. \tag{9.140}$$

Important solutions of (9.140) are:

$$\rho = \delta(H - E), \tag{9.141}$$

which is denoted by the **microcanonical ensemble**, where the total energy of the system is precisely known. If only the average (9.129) of the energy $< H >$ is known due to an interaction with a **heat bath**, ρ becomes

$$\rho = \exp\left(-H/(k_B T)\right), \tag{9.142}$$

which is denoted by the **canonical ensemble**. In (9.142) then T can be identified with the phenomenological temperature of the system while k_B is the **Boltzmann constant**.

Furthermore, if the particle number N is a constant of motion,

$$\{N, H\} = 0, \tag{9.143}$$

and only known on average $< N >$, ρ becomes

$$\rho = \exp\left(-\frac{H}{kT} - \alpha N\right), \tag{9.144}$$

where the Lagrange parameter α is related to the chemical potential and the temperature (see thermodynamics). Such an ensemble is called **grand-canonical**.

In addition to the microcanonical, canonical and grand-canonical ensembles the statistical physics also includes further ensembles, each of which is characterized by whether a thermodynamic observable is preserved **exactly** or **only on average**. These distinctions play no role for very large particle numbers in classical statistics but are of great importance for **quantum statistics**, where the **thermodynamic potentials**—similar to the generating functions $F_1, .., F_4$—emerge from each other by Legendre transformations (see thermodynamics).

In summarizing this chapter we have formulated the dynamics in phase-space variables, i.e. in generalized coordinates and generalized momenta. In this case the time evolution of an observable, that not explicitly depends on time, is given by **Poisson brackets**, which are determined by the derivative of the observable and the Hamiltonian with respect to the phase-space variables. The elementary Poisson bracket between generalized coordinates and generalized momenta was shown to be unity for associated pairs and their time evolution to be given by the Poisson bracket with the Hamilton function, i.e. the canonical equations of motion. The Poisson brackets thus allow for an algebraic formulation of the dynamics. Furthermore, we have proven that point transformations and extended canonical transformations between the generalized coordinates and momenta keep the equations of motion **canonical invariant**. Furthermore, the **elementary Poisson brackets** were proven to be invariant with respect to canonical transformations such that a formulation of classical mechanics could be achieved, that is independent on the choice of the generalized coordinates. This will pave the way to quantum mechanics, where the Poisson brackets will be replaced by commutators of operators in an abstract Hilbert space. The algebraic formulation also leads to a rigid formulation of statistical mechanics, where the physical system - in equilibrium - is described by ensembles with properties that are defined by expectation values of conserved quantities.

Appendix

Appendix

In these appendices some useful extensions are presented: The Lagrange and Hamilton functions for relativistic systems are given as well as for continuum mechanics. We will close by providing numerical algorithms for differentiation and integration as well as for the numerical solution of a set of differential equations.

A.1 Relativistic Mechanics

Using the example of the relativistic treatment of a charged particle in an electromagnetic field we want to show how the Lagrange and Hamilton formalism can be transferred and related to other areas of physics.

A.1.1 Lagrange Function for a Relativistic Particle

We are looking for a Lagrange function that leads to the equation of motion

$$\frac{d}{dt}(m(v)\overrightarrow{v}) = \overrightarrow{F} \tag{A.1}$$

with

$$m(v) = \gamma(v)m_0 = \frac{m_0}{\sqrt{1-v^2/c^2}} \tag{A.2}$$

and

$$\overrightarrow{F} = q(\overrightarrow{E} + (\overrightarrow{v} \times \overrightarrow{B})) \tag{A.3}$$

for the case of the Lorentz force. The fundamental relationships of Lagrangian mechanics,

$$p_i = \frac{\partial L}{\partial v_i}, \tag{A.4}$$

for the generalized momenta (A.4) as well as the Lagrange equations remain,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v_i}\right) = \frac{\partial L}{\partial x_i}. \tag{A.5}$$

Since in comparison to the non-relativistic case only (A.2) changes, it thus makes sense to start with:

$$\boxed{L = \tilde{T} - q\Phi + q\overrightarrow{v} \cdot \overrightarrow{A},} \tag{A.6}$$

where $\tilde{T}$ must be constructed in such a way that

$$\frac{\partial \tilde{T}}{\partial v_i} = m(v)v_i. \tag{A.7}$$

The solution is (up to a constant of integration)

$$\tilde{T} = -m_0c^2\sqrt{1 - v^2/c^2} = -\frac{m_0c^2}{\gamma(v)}, \tag{A.8}$$

obviously different from the kinetic energy

$$T = \frac{m_0c^2}{\sqrt{1-v^2/c^2}} - m_0c^2 = m_0c^2(\gamma(v) - 1). \tag{A.9}$$

By inserting (A.8), (A.6) into (A.5) we get–as desired–Eqs. (A.1)–(A.3).

A.1.2 Hamilton Function for a Relativistic Particle

The Hamilton function turns out to be identical to the energy:

$$H = \sum_{i=1}^{3} v_i p_i + m_0c^2\sqrt{1 - v^2/c^2} + q\Phi - q\sum_{i=1}^{3} v_i A_i =$$

$$\frac{m_0v^2}{\sqrt{1-v^2/c^2}} + m_0c^2\sqrt{1 - v^2/c^2} + q\Phi = T + q\Phi + m_0c^2 = E, \tag{A.10}$$

with

$$p_i = \frac{\partial L}{\partial v_i} = m(v)v_i + qA_i. \tag{A.11}$$

A.2 Continuum Mechanics

A.2.1 Lagrange Function for an Oscillating String

We assume a (long) linear chain of mass points (see Fig. A.1); their Lagrange function in case of harmonic forces (limited to **next neighbor interactions**) is:

$$L = \frac{m}{2}\sum_i \dot{q}_i^2 - \frac{k}{2}\sum_i (q_{i+1} - q_i)^2. \tag{A.12}$$

The generalized coordinates q_i are the deflections of the particles from the equilibrium position, $\dot{q}_i$ are the associated generalized velocities (Fig. A.1).

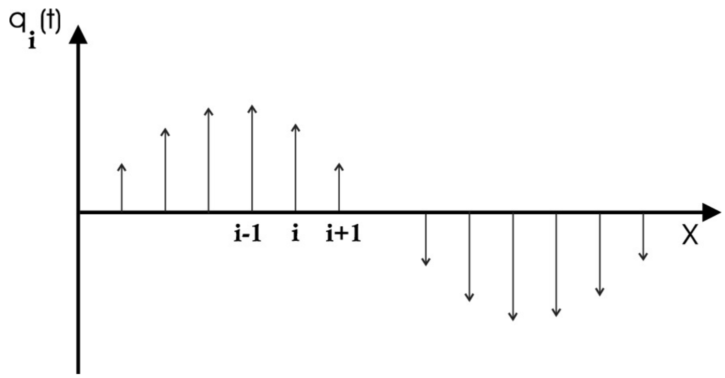

Fig. A.1 Illustration of an oscillating string for mass points at equal distances

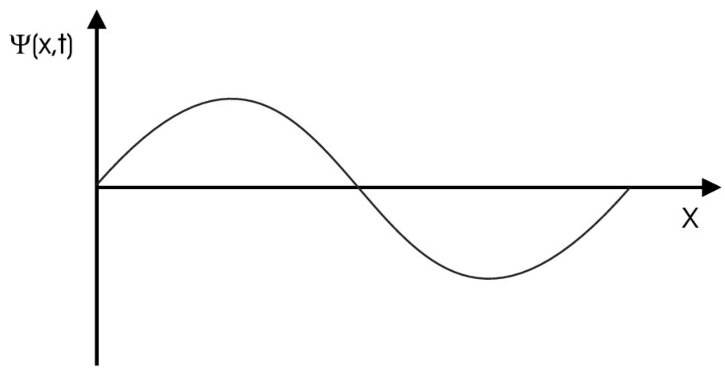

Fig. A.2 Illustration of an oscillating string in the continuum limit

The well-known equations of motion resulting from (A.12) are coupled harmonic oscillators:

$$m\ddot{q}_i - k(q_{i+1} - q_i) + k(q_i - q_{i-1}) = 0. \tag{A.13}$$

For the **transition to the continuum** (see Fig. A.2) let us reformulate (A.13) with $\mu = m/a$ and $\kappa = ka$ as:

$$L = \sum_i \left(\frac{\mu}{2}\dot{q}_i^2 - \kappa \frac{(q_{i+1} - q_i)^2}{2a^2} \right) a = \sum_i a L_i \tag{A.14}$$

and replace (in the limit $a \to 0$)

$$i \to x;\ \ \sum_i \ldots \to \int dx \ldots;\ q_i \to \psi(x;t);\ \dot{q}_i \to \frac{\partial \psi}{\partial t};\ \frac{1}{a}(q_{i+1} - q_i) \to \frac{\partial \psi}{\partial x}. \tag{A.15}$$

Then

$$\boxed{L = \frac{1}{2} \int \left(\mu \left(\frac{\partial \psi}{\partial t}\right)^2 - \kappa \left(\frac{\partial \psi}{\partial x}\right)^2 \right) dx = \int \mathscr{L}\ dx. \tag{A.16}}$$

Allowing that in general

(A.17)

$$\mathscr{L} = \mathscr{L}(\psi, \frac{\partial\psi}{\partial t}, \frac{\partial\psi}{\partial x};t),$$

the (generalized) Hamilton's variational principle,

$$\int\left(\int \mathscr{L}(\psi, \frac{\partial\psi}{\partial t}, \frac{\partial\psi}{\partial x};t)\ dx\right)\ dt = \text{extremum}, \tag{A.18}$$

leads to the associated Euler equations:

$$\frac{\partial}{\partial t}\left(\frac{\partial\mathscr{L}}{\partial(\frac{\partial\psi}{\partial t})}\right) + \frac{\partial}{\partial x}\left(\frac{\partial\mathscr{L}}{\partial(\frac{\partial\psi}{\partial x})}\right) = \frac{\partial\mathscr{L}}{\partial\psi}, \tag{A.19}$$

in analogy to

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{q}_i}\right) = \frac{\partial L}{\partial q_i}. \tag{A.20}$$

Especially in the case above (A.16) from (A.19) one gets the known vibration equation

$$\left(\frac{\partial\psi}{\partial x}\right)^2 - \frac{\mu}{\kappa}\left(\frac{\partial\psi}{\partial t}\right)^2 = 0. \tag{A.21}$$

A.2.2 Hamilton Function for an Oscillating String

Instead of the generalized momentum in the discrete case,

$$p_i = \frac{\partial L}{\partial\dot{q}_i}, \tag{A.22}$$

we have accordingly:

$$\pi(x,t) = \frac{\partial\mathscr{L}}{\partial(\frac{\partial\psi}{\partial t})}, \tag{A.23}$$

and can use the **Lagrange density** $\mathscr{L}$ to define the **Hamilton density** by a Legendre transformation

$$h = \pi\frac{\partial\psi}{\partial t} - \mathscr{L}. \tag{A.24}$$

The Hamilton function is then the spatial integral of h:

$$H = \int h \; dx = \int \left(\pi \frac{\partial \psi}{\partial t} - \mathscr{L} \right) \; dx \tag{A.25}$$

in analogy to the discrete case

$$H = \sum_i p_i \dot{q}_i - L. \tag{A.26}$$

Extensions:
(1.) The generalization to 3 spatial dimensions is simple:

$$x \rightarrow x_l; \qquad \int dx \ldots \rightarrow \int dx_1 dx_2 dx_3 \ldots . \tag{A.27}$$

and $(l = 1, 2, 3)$

$$\psi(x, t) \rightarrow \psi(x_l, t); \qquad \frac{\partial \psi}{\partial x} \rightarrow \frac{\partial \psi}{\partial x_l}. \tag{A.28}$$

(2.) In the case of electrodynamics not only a single field function occurs $\psi(\overrightarrow{r}, t)$, but 4 independent field functions forming a four-vector:

$$(A_\mu(\overrightarrow{r}, t)) = (\tfrac{i}{c} \Phi(\overrightarrow{r}, t), \overrightarrow{A}(\overrightarrow{r}, t)). \tag{A.29}$$

The general equations for the four-field $A_\mu(\overrightarrow{r}, t)$ with $(\mu = 0, 1, 2, 3)$ are the subject of **electrodynamics**.

A.3 Numerical Methods

Finally, we present the most important numerical algorithms for solving problems in mechanics.

A.3.1 Differentiation

Let a function $f_n = f(x_n)$ be defined on a grid with the same distance *h*, i.e.

$$f_n = f(x_n); x_n = nh; \quad (n = 0, \pm 1, \pm 2, \ldots). \tag{A.30}$$

To calculate the derivative of the function $f(x_n)$ at the position $x = 0$ we expand *f* in the neighborhood of *x* in a Taylor series

$$f(x) = f_0 + x f' + \frac{x^2}{2} f'' + \frac{x^3}{3!} f''' + \ldots \tag{A.31}$$

where all derivatives have to be calculated at the point $x = 0$. This gives the function *f* at the grid points $x_{\pm 1}$ via

$$f_{\pm 1} = f_0 \pm h f' + \frac{h^2}{2} f'' \pm \frac{h^3}{6} f''' + O(h^4). \tag{A.32}$$

By $O(h^4)$ terms of order h^4 or higher powers of *h* are summarized. Furthermore:

$$f_{\pm 2} = f_0 \pm 2\,hf' + \frac{4h^2}{2} f'' \pm \frac{8h^3}{6} f''' + O(h^4). \tag{A.33}$$

After subtracting f_{-1} from f_1 in (A.32) and rearranging the terms:

$$f' = \frac{f_1 - f_{-1}}{2h} - \frac{h^2}{6} f''' + O(h^4), \tag{A.34}$$

where the term $\sim f'''$ vanishes for sufficiently small *h*. The difference formula

$$f' = \frac{f_1 - f_{-1}}{2h} \tag{A.35}$$

is exact if the function *f* - in the interval $[-h, h]$—is a polynomial of second order because higher derivatives vanish.

By suitable combinations of (A.32), (A.33) difference formulae for higher derivatives can be specified. For example, one finds directly that

$$f_1 - 2f_0 + f_{-1} = h^2 f'' + O(h^4). \tag{A.36}$$

Then the second derivative of *f* at place $x = 0$ with a precision of order h^2 is

$$\frac{f_1 - 2f_0 + f_{-1}}{h^2} \approx f''. \tag{A.37}$$

For the 3rd derivative of *f* in $x = 0$ one obtains

$$\frac{f_2 - 2f_1 + 2f_{-1} - f_{-2}}{2\,h^3} \approx f'''. \tag{A.38}$$

Note: For calculating the derivative of *f* at position x_n one moves the arguments in the discrete formulae around *n*.

A.3.2 Integration

For the integration of a function *f*(*x*) in the interval [*a*, *b*] one divides the integral as:

$$\int_a^b f(x)dx = \int_a^{a+2h} f(x)dx + \int_{a+2h}^{a+4h} f(x)dx + \int_{a+4h}^{a+6h} f(x)dx + \ldots + \int_{b-2h}^{b} f(x)dx. \tag{A.39}$$

The underlying idea is now to replace the function *f* within the integration interval $[-h, h]$ by an approximate function (with the same values at the grid points), which can easily be integrated exactly. The simplest function is a linear approximation, which yields the **trapezoid formula**

$$\int_{-h}^{h} f(x)dx = \frac{h}{2}(f_{-1} + 2f_0 + f_1) + O(h^3). \tag{A.40}$$

More precise integration formulae can be found again by using the Taylor expansions (A.32), (A.33):

$$f(x) = f_0 + \frac{f_1 - f_{-1}}{2h} x + \frac{f_1 - 2f_0 + f_{-1}}{2h^2} x^2 + O(x^3). \tag{A.41}$$

This expression can be integrated elementary and we obtain the **Simpson rule**,

$$\int_{-h}^{h} f(x)dx = \frac{h}{3}(f_1 + 4f_0 + f_{-1}) + O(h^5), \tag{A.42}$$

which is more accurate than (A.40) by 2 orders in h. With (A.42) the integral (A.39) is approximated by:

$$\int_a^b f(x)dx = \frac{h}{3}[f(a) + 4f(a+h) + 2f(a+2h) + 4f(a+3h)$$

$$+2f(a+4h) + 4f(a+5h) + \ldots + 4f(b-h) + f(b)]. \tag{A.43}$$

Taking into account higher order terms in the Taylor expansion yields the **Bode formula**

$$\int_{x_0}^{x_4} f(x)dx = \frac{2h}{45}(7f_0 + 32f_1 + 12f_2 + 32f_3 + 7f_4) + O(h^7), \tag{A.44}$$

which is more precise by 2 orders in h than (A.42), but also clearly increases the computing effort.

A.3.3 Ordinary Differential Equations

The most general form of an ordinary differential equation is a set of $M = 2s$ coupled first-order equations,

$$\frac{d\mathbf{y}}{dt} = f(\mathbf{y}, t), \tag{A.45}$$

with an independent variable t and an M-dimensional vector $\mathbf{y} = (y_1, \ldots, y_M)$, such as the canonical equations of motion in Hamiltonian dynamics. The task is now to determe the value of $\mathbf{y}(t)$ if an initial value of $\mathbf{y}(t_0) = \mathbf{y}_0$ is given.

One of the simplest algorithms is the **Euler method**, in which the equation (A.45) is considered at the point t_n and the derivative on the left is replaced by the forward difference approximation:

$$\frac{\mathbf{y}_{n+1} - \mathbf{y}_n}{h} + O(h) = f(\mathbf{y}_n, t_n). \tag{A.46}$$

Then $\mathbf{y}_{n+1}$ can be calculated by a recursion formula from $\mathbf{y}_n$:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + hf(\mathbf{y}_n, t_n) + O(h^2). \tag{A.47}$$

This expression has a local error of order h^2 since the error of the forward difference formula is $O(h)$. The global error for N integration steps from $t = 0$ to $t = 1$ is then of order $NO(h^2) \approx O(h)$. This mistake only decreases linearly with the step size $h = \Delta t$.

Another way to solve the differential equation with higher accuracy is to set up recursion formulae in which $\mathbf{y}_{n+1}$ is not only linked with $\mathbf{y}_n$, but also with $\mathbf{y}_{n-1}, \mathbf{y}_{n-2}, \mathbf{y}_{n-3}, \ldots$. In order to derive such formulae explicitly, we integrate each single step of the differential equation exactly and get:

(A.48)

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \int_{t_n}^{t_{n+1}} f(\mathbf{y}, t)\ dt.$$

One can now use the values of $\mathbf{y}$ at t_n and t_{n-1} to find a linear extrapolation of *f* for the required integral:

$$f(\mathbf{y}, t) \approx \frac{t-t_{n-1}}{h} f(\mathbf{y}, t_n) - \frac{t-t_n}{h} f(\mathbf{y}, t_{n-1}) + O(h^2). \tag{A.49}$$

Inserting (A.49) into (A.48) and performing the *t*-integral, one obtains the **two-step method of Adams-Bashforth**:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h(\tfrac{3}{2} f_n - \tfrac{1}{2} f_{n-1}) + O(h^2). \tag{A.50}$$

Higher order methods can be achieved by using the *f* extrapolation with a higher order polynomial. When approximated by a cubic polynomial this results in the **four-step process of Adams and Bashforth**:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \tfrac{h}{24}(55 f_n - 59 f_{n-1} + 37 f_{n-2} - 9 f_{n-3}) + O(h^4). \tag{A.51}$$

For these algorithms the knowledge of the initial value alone is not sufficient to start the algorithms. That's why it is necessary to calculate the values of $\mathbf{y}$ at the first support points first, for example, by using the **Runge-Kutta** method.

The previous methods are **explicit**, since $\mathbf{y}_{n+1}$ is calculated from the known values of $\mathbf{y}_n$. **Implicit** algorithms, in which an equation must be solved, pave another way to achieve a higher accuracy. As an example let's mention the

Runge-Kutta algorithm of second order, which is often used. To this aim we approximate the function *f* in the integral of (A.48) by its Taylor expansion around the **middle point** of the integration interval and get

$$\mathbf{y}_{n+1} = \mathbf{y}_n + h f(\mathbf{y}_{n+1/2}, t_{n+1/2}) + O(h^3). \tag{A.52}$$

Since the error is of order $O(h^3)$ an approximation of $f(\mathbf{y}_{n+1/2}, t_{n+1/2})$ is of order $O(h^2)$ and good enough as provided by the simple Euler method (A.46). Defining now *k* as an intermediate approximation for the double difference between $\mathbf{y}_{n+1/2}$ and $\mathbf{y}_n$ then we can calculate $\mathbf{y}_{n+1}$ from $\mathbf{y}_n$ by the following two-step procedure:

$$k = h f(\mathbf{y}_n, t_n); \qquad \mathbf{y}_{n+1} = \mathbf{y}_n + h f(\mathbf{y}_n + \tfrac{k}{2}, t_n + \tfrac{h}{2}) + O(h^3). \tag{A.53}$$

The advantage of the Runge-Kutta method is that there are no special restrictions on the function *f* such as easy differentiability or linearity in $\mathbf{y}$. It also uses only the value of $\mathbf{y}$ at a single preceding point as opposed to the multi-step process described above. Equation (A.53), however, is required which implies that the value *f* at each integration step has to be calculated twice.

Runge-Kutta algorithms of higher order can be derived relatively directly. For this we use higher order integration formulae (see subchapter A.3.2) in order to replace the integral (A.48) by a finite sum of *f* values. For example, the Simpson rule gives:

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \tfrac{h}{6}\left[f(\mathbf{y}_n, t_n) + 4f(\mathbf{y}_{n+1/2}, t_{n+1/2}) + f(\mathbf{y}_{n+1}, t_{n+1})\right] + O(h^5). \tag{A.54}$$

The algorithm is completed by the fact that successive approximations are used for the $\mathbf{y}$'s (with a comparable accuracy) in the right side of (A.54). A **third-order algorithm** with a local error $O(h^4)$ then is:

$$k_1 = hf(\mathbf{y}_n, t_n);$$

$$k_2 = hf\left(\mathbf{y}_n + \tfrac{k_1}{2}, t_n + \tfrac{h}{2}\right);$$

$$k_3 = hf(\mathbf{y}_n - k_1 + 2k_2, t_n + h);$$

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \tfrac{1}{6}[k_1 + 4k_2 + k_3] + O(h^4). \tag{A.55}$$

It is based on Simpson's formula (A.42) and requires a threefold calculation of the function *f* per integration step.

Runge-Kutta algorithm of fourth order: By experience we have learned that a fourth-order algorithm, which needs 4 function *f* calculations per integration step, gives the best balance between accuracy and numerical effort. The algorithm for 4 intermediate variables k_i is:

$$k_1 = hf(\mathbf{y}_n, t_n);$$

$$k_2 = hf\left(\mathbf{y}_n + \tfrac{k_1}{2}, t_n + \tfrac{h}{2}\right);$$

$$k_3 = hf\left(\mathbf{y}_n + \tfrac{k_2}{2}, t_n + \tfrac{h}{2}\right);$$

$$k_4 = hf(\mathbf{y}_n + k_3, t_n + h);$$

$$\mathbf{y}_{n+1} = \mathbf{y}_n + \tfrac{1}{6}[k_1 + 2k_2 + 2k_3 + k_4] + O(h^5). \tag{A.56}$$

Index