Quantum Field Theory II: Quantum Electrodynamics

A Bridge between Mathematicians and Physicists

Bearbeitet von
Eberhard Zeidler

2009 2008. Buch. xxxvii, 1101 S. Hardcover
ISBN 978 3 540 85376 3
Format (B x L): 15,5 x 23,5 cm
Gewicht: 1841 g

1. Mathematical Principles of Modern Natural Philosophy

The book of nature is written in the language of mathematics.
Galileo Galilei (1564–1642)

At the beginning of the seventeenth century, two great philosophers, Francis Bacon (1561–1626) in England and René Descartes (1596–1650) in France, proclaimed the birth of modern science. Each of them described his vision of the future. Their visions were very different. Bacon said, "All depends on keeping the eye steadily fixed on the facts of nature." Descartes said, "I think, therefore I am." According to Bacon, scientists should travel over the earth collecting facts, until the accumulated facts reveal how Nature works. The scientists will then *induce* from the facts the laws that Nature obeys. According to Descartes, scientists should stay at home and *deduce* the laws of Nature by pure thought. In order to deduce the laws correctly, the scientists will need only the rules of logic and knowledge of the existence of God. For four hundred years since Bacon and Descartes led the way, science has raced ahead by following both paths simultaneously. Neither Baconian empiricism nor Cartesian dogmatism has the power to elucidate Nature's secrets by itself, but both together have been amazingly successful. For four hundred years English scientists have tended to be Baconian and French scientists Cartesian.
Faraday (1791–1867) and Darwin (1809–1882) and Rutherford (1871–1937) were Baconians; Pascal (1623–1662) and Laplace (1749–1827) and Poincaré (1854–1912) were Cartesians. Science was greatly enriched by the cross-fertilization of the two contrasting national cultures. Both cultures were always at work in both countries. Newton (1643–1727) was at heart a Cartesian, using pure thought as Descartes intended, and using it to demolish the Cartesian dogma of vortices. Marie Curie (1867–1934) was at heart a Baconian, boiling tons of crude uranium ore to demolish the dogma of the indestructibility of atoms.[1]

Freeman Dyson, 2004

It is important for him who wants to discover not to confine himself to a chapter of science, but keep in touch with various others.[2]

Jacques Hadamard (1865–1963)

[1] From Dyson's foreword to the book by P. Odifreddi, The Mathematical Century: The 30 Greatest Problems of the Last 100 Years, Princeton University Press, 2004. Reprinted by permission of Princeton University Press.

[2] J. Hadamard, The Psychology of Invention, Princeton University Press, 1945.

> Mathematics takes us still further from what is human, into the region of absolute necessity, to which not only the actual world, but every possible world must conform.[3]
>
> Bertrand Russel (1872–1972)

1.1 Basic Principles

There exist the following fundamental principles for the mathematical description of physical phenomena in nature.

(I) The infinitesimal principle due to Newton and Leibniz: The laws of nature become simple on an infinitesimal level of space and time.[4]

(II) The optimality principle (or the principle of least action): Physical processes proceed in such an optimal way that the action is minimal (or at least critical). Such processes are governed by ordinary or partial differential equations called the Euler–Lagrange equations.

(III) Emmy Noether's symmetry principle: Symmetries of the action functional imply conservation laws for the corresponding Euler–Lagrange equations (e.g., conservation of energy).

(IV) The gauge principle and Levi-Civita's parallel transport: The fundamental forces in nature (gravitational, eletromagnetic, strong, and weak interaction) are based on the symmetry of the action functional under local gauge transformations. The corresponding parallel transport of physical information generates the intrinsic Gauss–Riemann–Cartan–Ehresmann curvature which, roughly speaking, corresponds to the acting force (also called interaction). Briefly: force = curvature.

(V) Planck's quantization principle: Nature jumps.

(VI) Einstein's principle of special relativity: Physics does not depend on the choice of the inertial system.

(VII) Einstein's principle of general relativity: Physics does not depend on the choice of the local space-time coordinates of an observer.

(VIII) Dirac's unitarity principle: Quantum physics does not depend on the choice of the measurement device (i.e., on the choice of an orthonormal basis in the relevant Hilbert space). This corresponds to the invariance under unitary transformations.[5]

---

[3] The Earl of Russel was awarded the Nobel prize in literature in 1950. He worked in philosophy, mathematical logic, social sciences, and politics.

[4] I. Newton, Philosophiae naturalis principia mathematica (Mathematical principles of natural philosophy) (in Latin), 1687. Translated into English by A. Motte, in 1729, edited by F. Cajori, University of California Press, Berkeley, California, 1946. See also S. Chandrasekhar, Newton's Principia for the Common Reader, Oxford University Press, Oxford, 1997.

[5] Newton (1643–1727), Leibniz (1646–1716), Euler (1707–1783), Lagrange (1736–1813), Laplace (1749–1828), Legendre (1752–1833), Fourier (1768–1830), Gauss (1777–1855), Poisson (1781–1840), Faraday (1791–1867), Green (1793–1841),

Geometrization of physics. In mathematics, the properties of geometric objects do not depend on the choice of the coordinate system. This is similar to the principles (VI)–(VIII). Therefore, it is quite natural that geometric methods play a fundamental role in modern physics.

Linearity and nonlinearity. We have to distinguish between

(i) linear processes, and
(ii) nonlinear processes.

In case (i), the superposition principle holds, that is, the superposition of physical states yields again a physical state. Mathematically, such processes are described by linear spaces and linear operator equations. The mathematical analysis can be simplified by representing physical phenomena as superposition of simpler phenomena. This is the method of harmonic analysis (e.g., the Fourier method based on the Fourier series, the Fourier integral, or the Fourier–Stieltjes integral).

In case (ii), the superposition principle is violated. As a rule, interactions in nature are mathematically described by nonlinear operator equations (e.g., nonlinear differential or integral equations). The method of perturbation theory allows us to reduce (ii) to (i), by using an iterative method.

Basic properties of physical effects. For the mathematical investigation of physical effects, one has to take the following into account.

(A) Faraday's locality principle: Physical effects propagate locally in space and time (law of proximity theory).
(B) Green's locality principle: The response of a linear physical system can be described by localizing the external forces in space and time and by considering the superposition of the corresponding special responses (method of the Green's function). Furthermore, this can be used for computing nonlinear physical systems by iteration.
(C) Planck's constant: The smallest action (energy × time) in nature is given by the action quantum $h = 6.626\,0755 \cdot 10^{-34}$Js.
(D) Einstein's propagation principle: Physical effects travel at most with the speed of light c in a vacuum. Explicitly, $c = 2.997\,92458 \cdot 10^{8}$m/s.
(E) Gauge invariance principle: Physical effects are invariant under local gauge transformations. Physical experiments are only able to measure quantities which do not depend on the choice of the gauge condition.

Riemann (1826–1866), Maxwell (1831–1879), Lie (1842–1899), Klein (1849–1925), Poincaré (1854–1912), Planck (1858–1947), Élie Cartan (1859–1951), Hilbert (1862–1943), Minkowski (1864–1909), Levi-Civita (1873–1941), Einstein (1879–1955), Emmy Noether (1882–1935), Weyl (1885–1955), Schrödinger (1887–1961), Heisenberg (1901–1976), Dirac (1902–1984), Ehresmann (1905–1979), von Neumann (1903–1957), Tomonaga (1906–1979), Landau (1908–1968), Laurent Schwartz (1915–2002), Feynman (1918–1988), Schwinger (1918–1994), Yang (born 1922), Dyson (born 1923), Salam (1926–1996), Gell-Mann (born 1929), Glashow (born 1932), Weinberg (born 1933), Fritzsch (born 1943).

(F) The Planck scale hypothesis: Physics dramatically changes below the Planck length given by $l = 10^{-35}$m.

In what follows, let us discuss some basic ideas related to all of the principles summarized above. To begin with, concerning Faraday's locality principle, Maxwell emphasized the following:[6]

> Before I began the study of electricity I resolved to read no mathematics on the subject till I had first read through Faraday's 1832 paper *Experimental researches on electricity*. I was aware that there was supposed to be a difference between Faraday's way of conceiving phenomena and that of the mathematicians, so that neither he nor they were satisfied with each other's language. I had also the conviction that this discrepancy did not arise from either party being wrong. For instance, Faraday, in his mind, saw lines of force traversing all space where the mathematicians (e.g., Gauss) saw centers of force attracting at a distance; Faraday saw a medium where they saw nothing but distance; Faraday sought the seat of the phenomena in real actions going on in the medium, where they were satisfied that they had found it in a power of action at a distance impressed on the electric fluids.
>
> When I had translated what I considered to be Faraday's ideas into a mathematical form, I found that in general the results of the two methods coincide... I also found that several of the most fertile methods of research discovered by the mathematicians could be expressed much better in terms of the ideas derived from Faraday than in their original form.

1.2 The Infinitesimal Strategy and Differential Equations

> Differential equations are the foundation of the natural scientific, mathematical view of the world.
>
> Vladimir Igorevich Arnold (born 1937)

The infinitesimal strategy due to Newton and Leibniz studies the behavior of a physical system for infinitesimally small time intervals and infinitesimally small distances. This leads to the encoding of physical processes into differential equations (e.g., Newton's equations of motion in mechanics, or Maxwell's equations in electrodynamics).

> *The task of mathematics is to decode this information; that is, to solve the fundamental differential equations.*

1.3 The Optimality Principle

It is crucial that the class of possible differential equations is strongly restricted by the optimality principle. This principle tells us that the fundamental differential equations are the Euler–Lagrange equations to variational

[6] J. Maxwell, A Treatise on Electricity and Magnetism, Oxford University Press, Oxford, 1873.

problems. In 1918, Emmy Noether formulated her general symmetry principle in the calculus of variations. The famous Noether theorem combines Lie's theory of continuous groups with the calculus of variations due to Euler and Lagrange. This will be studied in Section 6.6.

1.4 The Basic Notion of Action in Physics and the Idea of Quantization

The most important physical quantity in physics is not the energy, but the action which has the physical dimension *energy times time*. The following is crucial.

(i) The fundamental processes in nature are governed by the principle of least action

$$\boxed{S = \min!}$$

where we have to add appropriate side conditions. In fact, one has to use the more general principal $S =$ critical! (principle of critical action). For example, if we consider the motion $q = q(t)$ of a particle of mass m on the real line, then the action is given by

$$S[q] := \int_{t_0}^{t_1} \left(\tfrac{1}{2} m\dot{q}(t)^2 - U(q(t))\right) dt.$$

Here, the function $U = U(q)$ is the potential, and the negative derivative, $-U'$, describes the acting force. In this case, the principle of critical action reads as

$$\boxed{S[q] = \text{critical!}, \qquad q(t_0) = q_0,\ q(t_1) = q_1} \tag{1.1}$$

where we fix the following quantities: the initial time t_0, the initial position q_0 of the particle, the final time t_1, and the final position q_1 of the particle. The solutions of (1.1) satisfy the Euler–Lagrange equation

$$m\ddot{q}(t) = F(t), \qquad t \in \mathbb{R}$$

with the force $F(t) = -U'(q(t))$. This coincides with the Newtonian equation of motion (see Sect. 6.5).

(ii) In 1900 Planck postulated that there do not exist arbitrarily small amounts of action in nature. The smallest amount of action in nature is equal to the Planck constant h. In ancient times, philosophers said:

Natura non facit saltus. (Nature does never make a jump.)

In his "Noveaux essais," Leibniz wrote:

Tout va par degrés dans la nature et rien par saut. (In nature everything proceeds little by little and not by jumping.)

In contrast to this classical philosophy, Planck formulated the hypothesis in 1900 that the energy levels of a harmonic oscillator form a discrete set. He used this fact in order to derive his radiation law for black bodies (see Sect. 2.3.1 of Vol. I). This was the birth of quantum physics. More generally, the energy levels of the bound states of an atom or a molecule are discrete. The corresponding energy jumps cause the spectra of atoms and molecules observed in physical experiments (e.g., the spectral analysis of the light coming from stars). Nowadays, we say that:

Nature jumps.

This reflects a dramatic change in our philosophical understanding of nature.

(iii) In order to mathematically describe quantum effects, one has to modify classical theories. This is called the process of quantization, which we will encounter in this series of monographs again and again. As an introduction to this, we recommend reading Chapter 7. Now to the point. Feynman discovered in the early 1940s in his dissertation in Princeton that the process of quantization can be most elegantly described by path integrals (also called functional integrals) of the form

$$\boxed{\int \mathrm{e}^{\mathrm{i}S[\psi]/\hbar}\,\mathcal{D}\psi}$$

where we sum over all classical fields ψ (with appropriate side conditions). Here, $\hbar := h/2\pi$. For example, the quantization of the classical particle considered in (i) can be based on the formula

$$G(q_0, t_0; q_1, t_1) = \int \mathrm{e}^{\mathrm{i}S[q]/\hbar}\,\mathcal{D}q.$$

Here, we sum over all classical motions $q = q(t)$ which satisfy the side condition $q(t_0) = q_0$ and $q(t_1) = q_1$. The Green's function G determines the time-evolution of the wave function ψ, that is, if we know the wave function $\psi = \psi(x, t_0)$ at the initial time t_0, then we know the wave function at the later time t by the formula

$$\psi(x,t) = \int_{\mathbb{R}} G(x,t;y,t_0)\psi(y,t_0)dy.$$

Finally, the wave function ψ tells us the probability

$$\int_a^b |\psi(x,t)|^2 dx$$

of finding the quantum particle in the interval $[a, b]$ at time t.

(iv) In quantum field theory, one uses the functional integral

$$\int \mathrm{e}^{\mathrm{i}S[\psi]/\hbar}\, \mathrm{e}^{\mathrm{i}\langle\psi|J\rangle}\, \mathcal{D}\psi$$

with the additional external source J. Differentiation with respect to J yields the moments of the quantum field. In turn, the moments determine the correlation functions (also called Green's functions). The correlation functions describe the correlations between different positions of the quantum field at different points in time. These correlations are the most important properties of the quantum field which can be related to physical measurements.

Feynman's functional integral approach to quantum physics clearly shows that both classical and quantum physics are governed by the classical action functional S. This approach can also be extended to the study of many-particle systems at finite temperature, as we have discussed in Sect. 13.8 of Vol. I. Summarizing, let us formulate the following general strategy:

The main task in modern physics is the mathematical description of the propagation of physical effects caused by interactions and their quantization.

In Sect. 1.9 we will show that in modern physics, interactions are described by gauge theories based on local symmetries.

1.5 The Method of the Green's Function

Basic ideas. As a prototype, consider the motion $x = x(t)$ of a particle of mass $m > 0$ on the real line under the action of the continuous force $F : \mathbb{R} \to \mathbb{R}$. The corresponding Newtonian equation of motion reads as

$$\boxed{m\ddot{x}(t) = F(t) \qquad \text{for all} \quad t \in \mathbb{R}} \tag{1.2}$$

with the initial condition

$$x(0) = a, \quad \dot{x}(0) = v.$$

We are given the initial position a and the initial velocity v at time $t = 0$. For simplifying notation, we set $m := 1$. In order to discuss Green's locality principle in physics, let us summarize the following facts. The unique solution of (1.2) reads as

$$x(t) = a + vt + \int_0^t (t - \tau)F(\tau)d\tau \qquad \text{for all} \quad t \in \mathbb{R}.$$

Equivalently, this can be written as

$$\boxed{x(t) = a + vt + \int_{-\infty}^{\infty} G(t,\tau)F(\tau)d\tau \qquad \text{for all} \quad t \in \mathbb{R}.} \tag{1.3}$$

The function G is called the Green's function of the differential equation (1.2). Explicitly,

$$G(t,\tau) := \begin{cases} t-\tau & \text{if} \quad 0 \le \tau \le t, \\ \tau - t & \text{if} \quad t \le \tau < 0, \\ 0 & \text{otherwise.} \end{cases}$$

Let us discuss the physical meaning of the Green's function G. To this end, for fixed positive number Δt and all times $t \in \mathbb{R}$, we introduce the Dirac Δt-delta function

$$\delta_{\Delta t}(t) := \begin{cases} \frac{1}{\Delta t} & \text{if} \quad 0 \le t \le \Delta t, \\ 0 & \text{otherwise.} \end{cases}$$

Obviously, we have

$$\lim_{\Delta t \to +0} \delta_{\Delta t}(t) = \begin{cases} +\infty & \text{if} \quad 0 \le t \le \Delta t, \\ 0 & \text{otherwise,} \end{cases}$$

and the normalization condition $\int_{-\infty}^{\infty} \delta_{\Delta t}(t)dt = 1$ is satisfied.

(i) Localized force. We are given the parameters $\Delta t > 0$ and $F_0 \in \mathbb{R}$. For fixed time t_0, we choose the special force

$$F(t) := F_0 \cdot \delta_{\Delta t}(t - t_0) \qquad \text{for all} \quad t \in \mathbb{R}.$$

By (1.3), the corresponding motion reads as

$$x_{\Delta t}(t) = a + vt + F_0 \cdot \frac{1}{\Delta t} \int_{t_0}^{t_0+\Delta t} G(t,\tau)d\tau \qquad \text{for all} \quad t \in \mathbb{R}.$$

Letting $\Delta t \to +0$, we get the motion[7]

$$x(t) = a + vt + F_0 \cdot G(t, t_0) \qquad \text{for all} \quad t \in \mathbb{R}. \tag{1.4}$$

This can be considered as the motion of the particle under the influence of the kick force $t \mapsto F_0\delta_{\Delta t}(t - t_0)$ at time t_0, as $\Delta t \to +0$. For $t_0 \ge 0$, the motion (1.4) looks like

[7] In fact, it follows from $\lim_{\varepsilon \to +0} G(t, t_0 + \varepsilon) = G(t, t_0)$ for all $t, t_0 \in \mathbb{R}$ that

$$\lim_{\Delta \to +0} \frac{1}{\Delta t} \int_{t_0}^{t_0+\Delta t} G(t,\tau)d\tau = G(t, t_0).$$

$$x(t) = \begin{cases} a + vt & \text{if} \quad t < t_0, \\ a + vt + F_0 \cdot (t - t_0) & \text{if} \quad t \geq t_0. \end{cases} \tag{1.5}$$

That is, the velocity jumps at time t_0. For $t_0 < 0$, (1.4) looks like

$$x(t) = \begin{cases} a + vt & \text{if} \quad t > t_0, \\ a + vt + F_0 \cdot (t_0 - t) & \text{if} \quad t \leq t_0. \end{cases} \tag{1.6}$$

(ii) Superposition of the original force by kick forces (physical interpretation of the Green's function). Fix $\Delta t > 0$. Consider the discrete points in time $n\Delta t$ where $n = 0, \pm 1, \pm 2, \ldots$ In terms of physics, let us approximate the given force $F = F(t)$ by a step function F_{approx}. That is, we use the superposition

$$F_{\text{approx}}(t) := \sum_{n=-\infty}^{\infty} F_n(t), \qquad t \in \mathbb{R}$$

of the kick forces $F_n(t) := F(n\Delta t)\delta_{\Delta t}(t - n\Delta t)\Delta t$. Explicitly,

$$F_n(t) = \begin{cases} F(n\Delta t) & \text{if} \quad n\Delta t \leq t \leq (n+1)\Delta t, \\ 0 & \text{otherwise.} \end{cases}$$

If Δt is sufficiently small, then the kick force F_n generates the approximate motion

$$x_n(t) = F(n\Delta t)G(t, n\Delta t)\Delta t, \qquad t \in \mathbb{R}$$

with $x_n(0) = 0$ and $\dot{x}_n(0) = 0$ for all $n = \pm 1, \pm 2, \ldots$ That is, the particle rests at the initial time $t = 0$. Consequently, by superposition, the force F_{approx} generates the approximate motion

$$x_{\text{approx}}(t) = \sum_{n=-\infty}^{\infty} x_n(t) = \sum_{n=-\infty}^{\infty} G(t, n\Delta t)F(n\Delta t)\Delta t, \qquad t \in \mathbb{R}.$$

As $\Delta t \to 0$, we get the motion $x(t) = \int_{-\infty}^{\infty} G(t, \tau)F(\tau)d\tau$ for all $t \in \mathbb{R}$.

The motions (1.5) and (1.6) have the following properties:

(a) $t \mapsto x(t)$ is continuous on $\mathbb{R}$.
(b) $t \mapsto x(t)$ is smooth on $\mathbb{R} \setminus \{t_0\}$, and $\ddot{x}(t) = 0$ for all $t \neq t_0$.
(c) $\dot{x}(t_0 + 0) = \dot{x}(t_0 - 0) + F_0$ (jump of the velocity at time t_0).
(d) In the sense of distributions, we have the following equation of motion:[8]

$$\boxed{m\ddot{x}(t) = F_0\delta(t - t_0), \qquad t \in \mathbb{R}.}$$

[8] We choose $m := 1$.

Let us prove (d). We have to show that

$$\int_{-\infty}^{\infty} x(t)\ddot{\varphi}(t)dt = F_0\varphi(t_0) \qquad \text{for all} \quad \varphi \in \mathcal{D}(\mathbb{R}).$$

Noting (b), integration by parts yields that $\int_{t_0}^{\infty} x(t)\ddot{\varphi}(t)dt$ is equal to

$$-x(t_0)\dot{\varphi}(t_0) - \int_{t_0}^{\infty} \dot{x}(t)\dot{\varphi}(t)dt = -x(t_0)\dot{\varphi}(t_0) + \dot{x}(t_0+0)\varphi(t_0).$$

Similarly, $\int_{-\infty}^{t_0} x(t)\ddot{\varphi}(t)dt = x(t_0)\dot{\varphi}(t_0) - \dot{x}(t_0-0)\varphi(t_0)$. Finally, use (c). □

Examples. Fix $t_0 := 0$ and $F_0 := 1$. If we choose $a = v := 0$, then the motion (1.5) looks like

$$x(t) = \theta(t)t \qquad \text{for all} \quad t \in \mathbb{R}.$$

Here, θ denotes the Heaviside function.[9] If we choose $a := 0$ and $v := -1$, then the motion (1.5) looks like

$$x(t) = -\theta(-t)t \qquad \text{for all} \quad t \in \mathbb{R}.$$

Finally, if we choose $a := 0, v := -\frac{1}{2}$, then the motion (1.5) looks like

$$x(t) = \tfrac{1}{2}(\theta(t)t - \theta(-t)t) = \tfrac{1}{2}|t| \qquad \text{for all} \quad t \in \mathbb{R}.$$

The relation between the theory of distributions and the method of averaging will be discussed in Sect. 1.7.

Iterative solution of nonlinear problems. The experience of physicists shows that

> *Interactions in nature lead to nonlinear terms in the corresponding differential equations.*

This explains the importance of nonlinear problems in physics. We want to show that the Green's function can also be used in order to investigate nonlinear problems. As a prototype, consider the differential equation

$$\boxed{m\ddot{x}(t) = -\kappa x(t)^3, \qquad t \in \mathbb{R}} \tag{1.7}$$

with the positive parameter κ called coupling constant, and the initial condition $x(0) = a, \dot{x}(0) = v$. This problem describes an anharmonic oscillator (see page 370). By (1.3), the initial-value problem (1.7) is equivalent to the nonlinear integral equation

$$\boxed{x(t) = a + vt - \kappa \int_{-\infty}^{\infty} G(t,\tau)x(\tau)^3 d\tau \qquad \text{for all} \quad t \in \mathbb{R},} \tag{1.8}$$

[9] Recall that $\theta(t) := 1$ if $t \geq 0$ and $\theta(t) := 0$ if $t < 0$.

by setting $F(t) := -\kappa x(t)^3$. The corresponding iterative method reads as

$$x_{n+1}(t) = a + vt - \kappa \int_{-\infty}^{\infty} G(t,\tau) x_n(\tau)^3 d\tau, \qquad n = -1, 0, 1, \ldots$$

with $x_{-1}(t) := 0$ for all $t \in \mathbb{R}$. This iterative method is also called the bootstrap method by physicists. In particular, $x_0(t) = a + vt$ for all $t \in \mathbb{R}$. The first approximation,

$$x_1(t) = a + vt - \kappa \int_{-\infty}^{\infty} G(t,\tau) x_0(\tau)^3 d\tau \qquad \text{for all} \quad t \in \mathbb{R}$$

is called the Born approximation by physicists. If the coupling constant κ is sufficiently small, then the iterative method converges to the solution of the original integral equation (1.8), that is, $\lim_{n\to\infty} x_n(t) = x(t)$ for all $t \in \mathbb{R}$.

> *The two problems* (1.7) *and* (1.8) *reflect a crucial duality between differential equations and integral equations. The kernel G of the integral equation* (1.8) *is the Green's function of the linearized differential equation* (1.2).

In this series of monographs, we will frequently use this duality. For example, in Sect. 8.6 we will study stationary scattering processes in quantum mechanics by replacing the Schrödinger differential equation by the dual Lippmann–Schwinger integral equation.

> *Therefore, nonlinear problems can be iteratively solved if the Green's function is known.*

This is the method of perturbation theory, which is basic for quantum field theory. For the computation of the Green's function, one can use Fourier's method. For the Newtonian motion (1.2), this will be studied in Sect. 2.2.14 in terms of the Fourier integral transform.

1.6 Harmonic Analysis and the Fourier Method

The superposition principle. In 1822 Fourier published his monograph *Théorie analytique de la chaleur* (analytic heat theory) where he used both the Fourier series and the Fourier integral in order to solve numerous problems in heat conduction. Let us sketch the basic ideas. For given time period $T > 0$, let us introduce the corresponding angular frequency

$$\Delta\omega := \frac{2\pi}{T}.$$

Fourier's method of harmonic analysis is the most important method for getting explicit solutions of linear partial differential equations in mathematical physics and for explicitly computing the corresponding Green's functions.[10]

[10] Much material can be found in P. Morse and H. Feshbach, Methods of Theoretical Physics, Vols. 1, 2, McGraw-Hill, New York, 1953. Fourier's method is intimately

(i) Discrete Fourier transformation. Let $f : \mathbb{R} \to \mathbb{C}$ be a smooth function of period $T > 0$. Then[11]

$$F(t) = \sum_{k=-\infty}^{\infty} a(k) \mathrm{e}^{\mathrm{i}tk\Delta\omega}, \qquad t \in \mathbb{R} \tag{1.9}$$

with the so-called Fourier coefficients

$$a(k) := \frac{1}{T} \int_{-T/2}^{T/2} F(t) \mathrm{e}^{-\mathrm{i}tk\Delta\omega} dt, \qquad k = 0, \pm 1, \pm 2, \ldots$$

Rigorously, the Fourier series (1.9) converges uniformly on the real line. Equation (1.9) tells us that the force function F can be represented by a superposition of special oscillating forces $t \mapsto a(k)\mathrm{e}^{\mathrm{i}tk\Delta\omega}$ of period T, angular frequency $k\Delta\omega$, and amplitude $a(k)$ with $k = 0, \pm 1, \pm 2, \ldots$ The map

$$F \mapsto \{a(k)\}_{k\in\mathbb{Z}}$$

is called the discrete Fourier transformation (with respect to the given period T).

(ii) Rescaling. Set $\hat{F}(k\Delta\omega) := Ta(-k)/\sqrt{2\pi}$ where $k = 0, \pm 1, \pm 2, \ldots$, and choose the angular frequencies

$$\omega := k\Delta\omega, \qquad k = 0, \pm 1, \pm 2, \ldots$$

Then

$$F(t) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \hat{F}(k\Delta\omega) \mathrm{e}^{-\mathrm{i}tk\Delta\omega} \Delta\omega, \qquad t \in \mathbb{R}$$

with

$$\hat{F}(\omega) := \frac{1}{\sqrt{2\pi}} \int_{-T/2}^{T/2} F(t) \mathrm{e}^{\mathrm{i}t\omega} dt, \qquad k = 0, \pm 1, \pm 2, \ldots$$

(iii) Continuous Fourier transformation. Suppose that the period T goes to infinity, $T \to \infty$. Then we formally obtain the integral

$$F(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{F}(\omega) \mathrm{e}^{-\mathrm{i}t\omega} d\omega, \qquad t \in \mathbb{R} \tag{1.10}$$

related to special functions in mathematical physics based on symmetries. We refer to N. Vilenkin and A. Klimyk, Special Functions and Representations of Lie Groups, Vols. 1–4, Kluwer, Dordrecht, 1991, and to A. Wawrzyńczyk, Group Representations and Special Functions, Reidel, Dordrecht, 1984 (see also the references for further reading about special functions in quantum mechanics on page 762).

[11] Recall that, by definition, $\sum_{k=-\infty}^{\infty} b(k) := \sum_{k=0}^{\infty} b(k) + \sum_{k=-1}^{-\infty} b(k)$.

with

$$\hat{F}(\omega) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(t) \mathrm{e}^{\mathrm{i}t\omega} dt, \qquad \omega \in \mathbb{R}. \tag{1.11}$$

Rigorously, if the function $F : \mathbb{R} \to \mathbb{C}$ is smooth and rapidly decreasing at infinity, that is, $F \in \mathcal{S}(\mathbb{R})$,[12] then the function F can be represented by (1.10) where the function $\hat{F}$ is given by (1.11). Moreover, we have $\hat{F} \in \mathcal{S}(\mathbb{R})$. Equation (1.10) tells us that each function $F \in \mathcal{S}(\mathbb{R})$ is the superposition of harmonic waves $t \mapsto \hat{F}(\omega)\mathrm{e}^{-\mathrm{i}t\omega}$ of angular frequency ω, and the corresponding amplitude function $\omega \mapsto \hat{F}(\omega)$ lies in $\mathcal{S}(\mathbb{R})$. The map $F \mapsto \hat{F}$ is called the continuous Fourier transformation (or, briefly, the Fourier transformation) from the time space to the frequency space.

Terminology. Passing from frequency ω to energy $E = \hbar\omega$, we define the Fourier transformation from the time space to the energy space by setting

$$F(t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \hat{F}_*(E) \mathrm{e}^{-\mathrm{i}Et/\hbar} dE, \qquad t \in \mathbb{R} \tag{1.12}$$

with

$$\hat{F}_*(E) := \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} F(t) \mathrm{e}^{\mathrm{i}Et/\hbar}\, dt, \qquad E \in \mathbb{R}. \tag{1.13}$$

Here, $\sqrt{\hbar}\hat{F}_*(\hbar\omega) = \hat{F}(\omega)$. This is also called the rescaled Fourier transformation. Motivated by the Fourier-Minkowski transformation in the 4-dimensional space-time (Minkowski space) in Einstein's theory of special relativity, we will distinguish between the Fourier transformation (1.12), (1.13) from the time space to the energy space and the Fourier transformation from the position space to the momentum space given by

$$F(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \hat{F}_{**}(p) \mathrm{e}^{\mathrm{i}xp/\hbar}\, dp, \qquad x \in \mathbb{R} \tag{1.14}$$

with

$$\hat{F}_{**}(p) := \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} F(x) \mathrm{e}^{-\mathrm{i}xp/\hbar}\, dx, \qquad p \in \mathbb{R}. \tag{1.15}$$

Note that $\sqrt{\hbar}\hat{F}_{**}(p) = \hat{F}(-p/\hbar)$ for all momenta $p \in \mathbb{R}$. This is discussed on page 538 of Vol. I. To simplify notation, we will frequently write $\hat{F}$ instead of $\hat{F}_*$ (resp. $\hat{F}_{**}$) if any misunderstanding is excluded.[13]

[12] The definition of the space $\mathcal{S}(\mathbb{R})$ can be found in Sect. 10.3.3 of Vol. I.

[13] In the literature, one also uses the asymmetrical formulas

$$F(x) = \int_{-\infty}^{\infty} \hat{F}_{\mathrm{asym}}(p) \mathrm{e}^{\mathrm{i}xp/\hbar}\, dp, \qquad x \in \mathbb{R}$$

Prototype of the Fourier method. Consider the differential equation

$$\boxed{\ddot{x}(t) = -\omega_0^2 x(t) + F(t), \qquad t \in \mathbb{R}.} \tag{1.16}$$

We are given the parameter $\omega_0 > 0$ and the periodic smooth force function $F : \mathbb{R} \to \mathbb{R}$ with the period $T > 0$. We are looking for a solution $x : \mathbb{R} \to \mathbb{R}$. In terms of physics, the function $x = x(t)$ describes the motion of a particle with mass $m = 1$ under the action of the external force $F(t)$ and the reactive force $-\omega_0^2 x(t)$ at time t. Physicists call this an harmonic oscillator. The angular frequency of the force F is given by $\Delta\omega := 2\pi/T$.

We postulate that $\omega_0 \neq k\Delta\omega$ for all integers k.

In terms of physics, this crucial condition means that the external force F is not in resonance with the eigenoscillations of the harmonic oscillator. The general solution of (1.16) reads as

$$\boxed{x(t) = \Re(a\mathrm{e}^{\mathrm{i}t\omega_0} + b\mathrm{e}^{-\mathrm{i}t\omega_0} + x_{\mathrm{special}}(t)), \qquad t \in \mathbb{R}} \tag{1.17}$$

with arbitrary complex numbers a, b. Furthermore, we have the special solution

$$x_{\mathrm{special}}(t) := \int_{-T/2}^{T/2} G(t,\tau)F(\tau)d\tau, \qquad t \in \mathbb{R},$$

and the Green's function

$$G(t,\tau) := \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \frac{\mathrm{e}^{\mathrm{i}(t-\tau)k\Delta\omega}}{\omega_0^2 - (k\Delta\omega)^2}\, \Delta\omega. \tag{1.18}$$

Let us prove this.

(I) Formal computation. In order to construct a special solution of (1.16), we start with the ansatz

$$x_{\mathrm{special}}(t) := \sum_{k=-\infty}^{\infty} b(k)\mathrm{e}^{\mathrm{i}tk\Delta\omega}.$$

Now insert this into (1.16) and use the Fourier series (1.9) for the force F. From $\ddot{x}(t) + \omega_0^2 x(t) - F(t) = 0$ we get

with

$$\hat{F}_{\mathrm{asym}}(p) = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} F(x)\mathrm{e}^{-\mathrm{i}xp/\hbar}\, dx, \qquad p \in \mathbb{R}.$$

However, whereas the transformations $F \mapsto \hat{F}$ and $F \mapsto \hat{F}_*$ generate unitary operators on the Hilbert space $L_2(\mathbb{R})$, this is not the case for $F \mapsto \hat{F}_{\mathrm{asym}}$. Therefore, the choice of the transformation $F \mapsto \hat{F}_{\mathrm{asym}}$ has the disadvantage that it violates the fundamental unitary symmetry between position and momentum in quantum physics.

$$\sum_{k=-\infty}^{\infty} \left(-(k\Delta\omega)^2 b(k) + \omega_0^2 b(k) - a(k)\right) \cdot \mathrm{e}^{\mathrm{i}tk\Delta\omega} = 0.$$

Hence $-(k\Delta\omega)^2 b(k) + \omega_0^2 b(k) - a(k) = 0$. This implies

$$x_{\text{special}}(t) = \sum_{k=-\infty}^{\infty} \frac{a(k)}{\omega_0^2 - (k\Delta\omega)^2} \cdot \mathrm{e}^{\mathrm{i}tk\Delta\omega}. \tag{1.19}$$

Noting that $a(k) = \frac{\Delta\omega}{2\pi} \int_{-T/2}^{T/2} F(\tau)\mathrm{e}^{-\mathrm{i}\tau k\Delta\omega} d\tau$, we get the desired formula (1.17), by formally interchanging summation and integration.

(II) Rigorous proof. Let $N = 1, 2, \ldots$ Since the function F is smooth, we have

$$a(k) = O\left(\frac{1}{k^N}\right) \qquad \text{for all} \quad k \in \mathbb{Z}.$$

The same is true for $b(k)$. By the majorant criterion, all the Fourier series involving $a(k)$ and $b(k)$ converge uniformly on the real line, and hence term-by-term differentiation (resp. integration) is allowed. This shows that the function x_{special} given by (1.19) is indeed a special solution of the inhomogeneous differential equation (1.16). Finally, note that the general solution of the homogeneous equation (1.16) with $F = 0$ is given by $x(t) = a\mathrm{e}^{\mathrm{i}t\omega_0} + b\mathrm{e}^{-\mathrm{i}t\omega_0}$ with arbitrary complex numbers a and b. This finishes the proof. □

Resonances and the singularities of the Green's function. Suppose that $\omega_0 = k_0\Delta\omega$ for some nonzero integer k_0. Then it follows from (1.18) that the Green's function G has a singularity if we choose $k = k_0$. In the case where the function F satisfies the condition $a(k_0) \neq 0$, that is,

$$\int_{-T/2}^{T/2} F(\tau)\mathrm{e}^{-\mathrm{i}\tau\omega_0} d\tau \neq 0,$$

physicists say that the external force F is in resonance with the eigenfrequency ω_0 of the harmonic oscillator.

Resonance effects cause singularities of the Green's function.

In the present case, the difficulty disappears if we demand that $a(k_0) = 0$. Then the singularity drops out in (1.19).

Resonances are responsible for complicated physical effects.

For example, the observed chaotic motion of some asteroids is due to resonance effects in celestial mechanics (the Kolmogorov–Arnold–Moser theory). In quantum field theory, internal resonances of the quantum field cause special quantum effects (e.g., the Lamb shift in the spectrum of the hydrogen atom and the anomalous magnetic moment of the electron), which have to be treated with the methods of renormalization theory (see Chap. 17 on radiative corrections in quantum electrodynamics).

1.7 The Method of Averaging and the Theory of Distributions

In the early 20th century, mathematicians and physicists noticed that for wave problems, the Green's functions possess strong singularities such that the solution formulas of the type (1.3) fail to exist as classical integrals.[14] In his classic monograph

The Principles of Quantum Mechanics,

Clarendon Press, Oxford, 1930, Dirac introduced a singular object $\delta(t)$ (the Dirac delta function), which is very useful for the description of quantum processes and the computation of Green's functions. In the 1940s, Laurent Schwartz gave all these approaches a sound basis by introducing the notion of distribution (generalized function). In order to explain Laurent Schwartz's basic idea of averaging, consider the continuous motion

$$x(t) := |t| \qquad \text{for all} \quad t \in \mathbb{R}$$

of a particle on the real line. We want to compute the force $F(t) = m\ddot{x}(t)$ acting on the particle at time t. Classically, $F(t) = 0$ if $t \neq 0$, and the force does not exist at the point in time $t = 0$. We want to motivate that

$$\boxed{F(t) = 2m\delta(t) \qquad \text{for all} \quad t \in \mathbb{R}.} \tag{1.20}$$

(I) The language of Dirac. For the velocity, $\dot{x}(t) = 1$ if $t > 0$, and $\dot{x}(t) = -1$ if $t < 0$. For $t = 0$, the derivative $\dot{x}(0)$ does not exist. We define $\dot{x}(0) := 0$. Hence

$$\dot{x}(t) = \theta(t) - \theta(-t).$$

Since $\dot{\theta}(t) = \delta(t)$, we get

$$\ddot{x}(t) = \delta(t) + \delta(-t) = 2\delta(t) \qquad \text{for all} \quad t \in \mathbb{R}.$$

Formally, $\delta(t) = 0$ if $t \neq 0$, and $\delta(0) = \infty$ with $\int_{-\infty}^{\infty} \delta(t)dt = 1$. Obviously, there is *no classical function* δ which has such properties.[15]

(II) The language of Laurent Schwartz. Choose $\varepsilon > 0$. We first pass to the regularized motion $x = x_\varepsilon(t)$ for all $t \in \mathbb{R}$. That is, the function $x_\varepsilon : \mathbb{R} \to \mathbb{R}$ is smooth for all $\varepsilon > 0$ and

[14] For example, see J. Hadamard, The Initial-Value Problem for Linear Hyperbolic Partial Differential Equations, Hermann, Paris (in French). A modern version of Hadamard's theory can be found in P. Günther, Huygens' Principle and Hyperbolic Differential Equations, Academic Press, San Diego, 1988. See also C. Bär, N. Ginoux, and F. Pfäffle, Wave Equations on Lorentzian Manifolds and Quantization, European Mathematical Society 2007.

[15] See the detailed discussion of the formal Dirac calculus in Sect. 11.2 of Vol. I.

$$\lim_{\varepsilon\to+0} x_\varepsilon(t) = |t| \qquad \text{for all} \quad t \in \mathbb{R},$$

where this convergence is uniform on all compact time intervals.[16] We introduce the averaged force

$$\mathsf{F}_\varepsilon(\varphi) := \int_{-\infty}^{\infty} m\ddot{x}_\varepsilon(t)\varphi(t)dt$$

for all averaging functions $\varphi \in \mathcal{D}(\mathbb{R})$ (i.e., $\varphi : \mathbb{R} \to \mathbb{C}$ is smooth and vanishes outside some bounded interval. In other words, φ has compact support.) Since x_ε is smooth, integration by parts twice yields

$$\mathsf{F}_\varepsilon(\varphi) = \int_{-\infty}^{\infty} mx_\varepsilon(t)\ddot{\varphi}(t)dt.$$

Letting $\varepsilon \to +0$, we define the mean force by

$$\mathsf{F}(\varphi) := \lim_{\varepsilon\to+0} \mathsf{F}_\varepsilon(\varphi) = \int_{-\infty}^{\infty} mx(t)\ddot{\varphi}(t)dt.$$

Integration by parts yields $\int_0^\infty |t|\ \ddot{\varphi}(t)dt = -\int_0^\infty \dot{\varphi}(t)dt = \varphi(0)$. Similarly, $\int_{-\infty}^0 |t|\ddot{\varphi}(t)dt = -\int_{-\infty}^0 \dot{\varphi}(t)dt = \varphi(0)$. Summarizing, we obtain the averaged force

$$\boxed{\mathsf{F}(\varphi) = 2m\varphi(0) \qquad \text{for all} \quad \varphi \in \mathcal{D}(\mathbb{R}).} \tag{1.21}$$

In the language of distributions, we have $\mathsf{F} = 2m\delta$, where δ denotes the Dirac delta distribution. A detailed study of the theory of distributions and its applications to physics can be found in Chaps. 11 and 12 of Vol. I. In particular, equation (1.20) is equivalent to (1.21), in the sense of distribution theory.

In terms of experimental physics, distributions correspond to the fact that measurement devices only measure averaged values. It turns out that classical functions can also be regarded as distributions. However, in contrast to classical functions, the following is true:

Distributions possess derivatives of all orders.

Therefore, the theory of distributions is the quite natural completion of the infinitesimal strategy due to Newton and Leibniz, who lived almost three hundred years before Laurent Schwartz. This shows convincingly that the development of mathematics needs time.

[16] For example, choose $x_\varepsilon(t) := r_\varepsilon(t)|t|$ for all $t \in \mathbb{R}$, where the regularizing function $r_\varepsilon : \mathbb{R} \to [0,1]$ is smooth, and $r_\varepsilon(t) := 1$ if $t \notin [-2\varepsilon, 2\varepsilon]$, as well as $r_\varepsilon(t) := 0$ if $t \in [-\varepsilon, \varepsilon]$.

1.8 The Symbolic Method

The symbol of the operator $\frac{d}{dt}$. Consider again the Fourier transformation

$$F(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{F}(\omega) \mathrm{e}^{-\mathrm{i}t\omega} d\omega, \qquad t \in \mathbb{R} \tag{1.22}$$

with

$$\hat{F}(\omega) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(t) \mathrm{e}^{\mathrm{i}t\omega t} dt, \qquad \omega \in \mathbb{R}. \tag{1.23}$$

Let $n = 1, 2, \ldots$ Differentiation of (1.22) yields

$$\frac{d^n}{dt^n} F(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (-\mathrm{i}\omega)^n \hat{F}(\omega) \mathrm{e}^{\mathrm{i}t\omega} d\omega, \qquad t \in \mathbb{R}. \tag{1.24}$$

The function

$$s(\omega) := -\mathrm{i}\omega \qquad \text{for all} \quad \omega \in \mathbb{R}$$

is called the *symbol* of the differential operator $\frac{d}{dt}$. For $n = 0, 1, 2, \ldots$, we have

$$\frac{d^n}{dt^n} F \Rightarrow s^n \hat{F}.$$

This means that the action of the differential operator $\frac{d^n}{dt^n}$, with respect to time t, can be described by the multiplication of the Fourier transform $\hat{F}$ by s^n in the frequency space.

This corresponds to a convenient algebraization of derivatives.

Over the centuries, mathematicians and physicists tried to simplify computations. The relation

$$\ln(ab) = \ln a + \ln b \qquad \text{for all} \quad a, b > 0 \tag{1.25}$$

allows us to reduce multiplication to addition. This fact was extensively used by Kepler (1571–1630) in order to simplify his enormous computations in celestial mechanics.

Similarly, the Fourier transformation allows us to reduce differentiation to multiplication.

Furthermore, there exists a natural generalization of the logarithmic function to Lie groups. Then the crucial formula (1.25) passes over to the transformation formula from the Lie group $\mathcal{G}$ to its Lie algebra $\mathcal{LG}$. This transformation is well defined for the group elements near the unit element (see Vol. III).

Pseudo-differential operators and Fourier integral operators. The modern theory of pseudo-differential operators (e.g., differential and integral

operators) and Fourier integral operators is based on the use of symbols of the form

$$s = s(\omega, t, \tau),$$

which depend on frequency ω, time t, and time τ. The expressions

$$(AF)(t) := \frac{1}{2\pi} \int_{\mathbb{R}^2} s(\omega, t, \tau) F(\tau) \mathrm{e}^{\mathrm{i}\omega(\tau - t)} \, d\tau d\omega, \quad t \in \mathbb{R} \tag{1.26}$$

and

$$(BF)(t) := \frac{1}{2\pi} \int_{\mathbb{R}^2} s(\omega, t, \tau) F(\tau) \mathrm{e}^{\mathrm{i}\varphi(\omega, t, \tau)} \, d\tau d\omega, \quad t \in \mathbb{R} \tag{1.27}$$

correspond to the pseudo-differential operator A and the Fourier integral operator B. If we choose the special phase function

$$\varphi(\omega, t, \tau) := \omega(t - \tau),$$

then the operator B passes over to A. If, in addition, the symbol s does not depend on τ, then integration over τ yields

$$(AF)(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} s(\omega, t) \hat{F}(\omega) \mathrm{e}^{-\mathrm{i}\omega t} \, d\omega.$$

In the special case where the symbol $s(\omega, t)$ only depends on the frequency ω, the pseudo-differential operator corresponds to the multiplication operator $\omega \mapsto s(\omega)\hat{F}(\omega)$ in the frequency space (also called Fourier space).

Long before the foundation of the theory of pseudo-differential operators and Fourier integral operators in the 1960s and 1970s, mathematicians and physicists used integral expressions of the form (1.26) and (1.27) in order to compute explicit solutions in electrodynamics (e.g., the Heaviside calculus and the Laplace transform applied to the study of electric circuits[17]), elasticity (singular integral equations), geometric optics (e.g., diffraction of light), and quantum mechanics.

> *The point is that the symbols know a lot about the properties of the corresponding operators, and an elegant algebraic calculus for operators can be based on algebraic operations for the symbols.*

As an introduction, we recommend:

> Yu. Egorov and M. Shubin, Foundations of the Classical Theory of Partial Differential Equations, Springer, New York, 1998 (Encyclopedia of Mathematical Sciences).
>
> Yu. Egorov, A. Komech, and M. Shubin, Elements of the Modern Theory of Partial Differential Equations, Springer, New York, 1999 (Encyclopedia of Mathematical Sciences).

[17] See E. Zeidler (Ed.), Oxford Users' Guide to Mathematics, Sect. 1.11, Oxford University Press, 2004.

F. Berezin and M. Shubin, The Schrödinger Equation, Kluwer, Dordrecht, 1991.

L. Faddeev and A. Slavnov, Gauge Fields, Benjamin, Reading, Massachusetts, 1980 (gauge theory, Weyl calculus, the Feynman path integral, and the Faddeev–Popov ghost approach to the Standard Model in particle physics).

We also refer to the following treatises:

L. Hörmander, The Analysis of Linear Partial Differential Operators. Vol. 1: Distribution Theory and Fourier Analysis, Vol. 2: Differential Operators with Constant Coefficients, Vol. 3: Pseudo-Differential Operators, Vol. 4: Fourier Integral Operators, Springer, New York, 1993.

R. Dautray and J. Lions, Mathematical Analysis and Numerical Methods for Science and Technology, Vols. 1–6, Springer, New York, 1988.

Heaviside's formal approach. Consider the differential equation

$$\frac{d}{dt}x(t) - x(t) = f(t). \tag{1.28}$$

We want to discuss the beauty, but also the shortcomings of the symbolic method due to Heaviside (1850–1925). Formally, we get

$$\left(\frac{d}{dt} - 1\right) x(t) = f(t).$$

Hence

$$x(t) = \frac{f(t)}{\frac{d}{dt} - 1}.$$

For complex numbers z with $|z| < 1$, we have the convergent geometric series $\frac{1}{z-1} = -1 - z - z^2 - z^3 + \ldots$ This motivates

$$x(t) = \left(-1 - \frac{d}{dt} - \frac{d^2}{dt^2} - \ldots\right) f(t). \tag{1.29}$$

If we choose $f(t) := t^2$, then

$$x(t) = -t^2 - 2t - 2. \tag{1.30}$$

Surprisingly enough, we get $\dot{x}(t) = -2t-2 = x(t)+t^2$. Therefore, the function $x(t)$ from (1.30) is a solution of (1.28). The same is true for all polynomials. To prove this, let f be a polynomial of degree $n = 0, 1, 2 \ldots$ Set

$$x(t) := -\sum_{k=0}^{n} \frac{d^k}{dt^k} f(t).$$

Then we get $\dot{x}(t) = -\sum_{k=0}^{n} \frac{d^{k+1}}{dt^{k+1}} f(t) = f(t) + x(t)$, since the $(n+1)$th derivative of f vanishes. However, the method above fails if we apply it to the exponential function $f(t) := \mathrm{e}^t$. Then

$$x(t) = \mathrm{e}^t + \mathrm{e}^t + \mathrm{e}^t + \dots,$$

which is meaningless. There arises the problem of establishing a more powerful method. In the history of mathematics and physics, formal (also called symbolic) methods were rigorously justified by using the following tools:

- the Fourier transformation,
- the Laplace transformation (which can be reduced to the Fourier transformation),
- Mikusiński's operational calculus based on the quotient field over a convolution algebra,
- von Neumann's operator calculus in Hilbert space,
- the theory of distributions,
- pseudo-differential operators and distributions (e.g., the Weyl calculus in quantum mechanics), and
- Fourier integral operators and distributions.

Mikusiński's elegant approach will be considered in Sect. 4.2 on page 191.

Motivation of the Laplace transformation via Fourier transformation. Consider the motion $x = x(t)$ of a particle on the real line with $x(t) = 0$ for all $t \le 0$. Suppose that the function $x : \mathbb{R} \to \mathbb{R}$ is continuous and bounded. The Fourier transform from the time space to the energy space reads as

$$\hat{x}(E) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} x(t)\mathrm{e}^{\mathrm{i}Et/\hbar}\, dt = \frac{1}{\sqrt{2\pi\hbar}} \int_{0}^{\infty} x(t)\mathrm{e}^{\mathrm{i}Et/\hbar}\, dt.$$

As a rule, this integral does not exist. To improve the situation, we fix the regularization parameter $\varepsilon > 0$, and we define the damped motion

$$x_\varepsilon(t) := x(t)\mathrm{e}^{-\varepsilon t} \qquad \text{for all} \quad t \in \mathbb{R}.$$

This is also called the adiabatic regularization of the original motion. Obviously, $\lim_{\varepsilon \to +0} x_\varepsilon(t) = x(t)$ for all $t \in \mathbb{R}$. The Fourier transform looks like

$$\hat{x}_\varepsilon(E) = \frac{1}{\sqrt{2\pi\hbar}} \int_{0}^{\infty} x(t)\mathrm{e}^{-\varepsilon t}\mathrm{e}^{\mathrm{i}Et/\hbar}\, dt = \frac{1}{\sqrt{2\pi\hbar}} \int_{0}^{\infty} x(t)\mathrm{e}^{\mathrm{i}\mathcal{E}t/\hbar}\, dt,$$

by introducing the complex energy $\mathcal{E} := E + \mathrm{i}\varepsilon$. To simplify notation, we set $\hbar := 1$.

Complex energies, damped oscillations, and the Laplace transform. The formal Heaviside calculus was justified by Doetsch in the 1930s by using the Laplace transform.[18] As a simple example, let us use the Laplace transformation in order to solve the differential equation (1.28). In particular, we will consider the case

[18] G. Doetsch, Theory and Applications of the Laplace Transform, Springer, Berlin, 1937 (in German). See also D. Widder, The Laplace Transform, Princeton University Press, 1944.

$$f(t) := \mathrm{e}^t$$

where the Heaviside method above fails. Let $x : [0, \infty[\to \mathbb{R}$ be a smooth function with the growth condition

$$|x(t)| \leq \text{const} \cdot \mathrm{e}^{\gamma_1 t} \qquad \text{for all} \quad t \geq 0$$

and fixed real number γ_1. The Laplace transform reads as

$$\boxed{\mathcal{L}(x)(\mathcal{E}) := \int_0^\infty x(t)\, \mathrm{e}^{\mathrm{i}\mathcal{E}t}\, dt, \qquad \Im(\mathcal{E}) > \gamma_1} \tag{1.31}$$

with the inverse transform

$$x(t) = \frac{1}{2\pi} PV \int_L (\mathcal{L}x)(\mathcal{E}) \mathrm{e}^{-\mathrm{i}\mathcal{E}t} d\mathcal{E}, \qquad t > 0 \tag{1.32}$$

on the real line $L := \{E + (\gamma_1 + 1)\mathrm{i} : \ E \in \mathbb{R}\}$ of the complex energy space. Here, we choose a system of units with $\hbar = h/2\pi := 1$ for Planck's action quantum.[19] The Laplace transform sends the function $t \mapsto x(t)$ on the time space to the function $\mathcal{E} \mapsto (\mathcal{L}x)(\mathcal{E})$ on the complex energy space. Here, it is crucial to use complex energies $\mathcal{E} = E - \Gamma\mathrm{i}$. In what follows, we will use the standard properties of the Laplace transformation which are proved in Sect. 2.2.6 of Vol. I. Let us start with an example. Choose the complex energy $\mathcal{E}_0 := E_0 - \Gamma_0\mathrm{i}$ with real values E_0 and Γ_0, and set[20]

$$\boxed{x(t) := \mathrm{e}^{-\mathrm{i}\mathcal{E}_0 t} = \mathrm{e}^{-\mathrm{i}E_0 t} \cdot \mathrm{e}^{-\Gamma_0 t}, \qquad t \in \mathbb{R}.} \tag{1.33}$$

Then, $\gamma_1 = -\Gamma_0 = \Im(\mathcal{E}_0)$, and we get

$$(\mathcal{L}x)(\mathcal{E}) = \frac{\mathrm{i}}{\mathcal{E} - \mathcal{E}_0}, \qquad \Im(\mathcal{E}) > \Im(\mathcal{E}_0).$$

Now to the point. We are given the smooth function $f : [0, \infty[\to \mathbb{C}$ with the growth condition

$$|f(t)| \leq \text{const} \cdot \mathrm{e}^{\gamma_0 t} \qquad \text{for all} \quad t \geq 0.$$

In order to solve the differential equation (1.28), we proceed as follows.

(I) Suppose first that the differential equation (1.28) has a smooth solution $x : [0, \infty[\to \mathbb{C}$ with $|x(t)| \leq \text{const} \cdot \mathrm{e}^{\gamma_1 t}$ for all $t \geq 0$ with $\gamma_1 \geq \gamma_0$. Then

[19] Set $\gamma_2 := \gamma_1 + 1$. The principal value of the integral is defined by

$$PV \int_L g(\mathcal{E}) d\mathcal{E} := \lim_{E_0 \to +\infty} \int_{-E^0 + \gamma_2 \mathrm{i}}^{E_0 + \gamma_2 \mathrm{i}} g(E + \gamma_2 \mathrm{i}) dE.$$

[20] If $E_0 > 0$ and $\Gamma_0 > 0$ then (1.33) is a damped oscillation with angular frequency $\omega_0 := E_0/\hbar = E_0$ and mean lifetime $\Delta t = \Gamma_0/\hbar = \Gamma_0$.

the Laplace transforms $\mathcal{L}x$ and $\mathcal{L}f$ exist for all $\mathcal{E} \in \mathbb{C}$ with $\Im(\mathcal{E}) > \gamma_1$. Furthermore,

$$(\mathcal{L}\dot{x})(\mathcal{E}) = -\mathrm{i}\mathcal{E}(\mathcal{L})(\mathcal{E}) - x(+0),$$

that is, the Laplace transforms converts differentiation into multiplication and translation in the complex energy space. By (1.28),

$$-\mathrm{i}\mathcal{E}(\mathcal{L}x)(\mathcal{E}) - (\mathcal{L}x)(\mathcal{E}) - x(+0) = (\mathcal{L}f)(\mathcal{E}), \qquad \Im(\mathcal{E}) > \gamma_1.$$

This yields the Laplace transform of the solution $t \to x(t)$, namely,

$$(\mathcal{L}x)(\mathcal{E}) = \frac{\mathrm{i}x(+0)}{\mathcal{E} - \mathrm{i}} + \frac{\mathrm{i}(\mathcal{L}f)(\mathcal{E})}{\mathcal{E} - \mathrm{i}}.$$

Setting $g(t) := \mathrm{e}^t$, we get $(\mathcal{L}g)(\mathcal{E}) = \frac{\mathrm{i}}{\mathcal{E}-\mathrm{i}}$. Therefore,

$$\mathcal{L}x = (\mathcal{L}g)x(+0) + (\mathcal{L}g)(\mathcal{L}f).$$

The convolution rule from Sect. 2.2.6 of Vol. I tells us that

$$x = gx(+0) + g * f.$$

Explicitly, this reads as

$$\boxed{x(t) = \mathrm{e}^t x(+0) + \int_0^t \mathrm{e}^{(t-\tau)} f(\tau) d\tau.} \tag{1.34}$$

Our argument shows that a solution of (1.28) has necessarily the form (1.34).

(II) Conversely, differentiation yields

$$\dot{x}(t) = \mathrm{e}^t x(+0) + f(t) + \int_0^t \mathrm{e}^{(t-\tau)} f(\tau) d\tau = x(t) + f(t)$$

for all $t \geq 0$. Consequently, the function $x = x(t)$ given by (1.34) is indeed a solution of the original differential equation (1.28) for all times $t \geq 0$. For example, if $f(t) := \mathrm{e}^t$, then

$$x(t) = \mathrm{e}^t x(+0) + t\mathrm{e}^t.$$

This is a solution of (1.28) for all times $t \in \mathbb{R}$.

The same method of the Laplace transformation can be applied to general systems of ordinary differential equations with constant coefficients. Such equations are basic for the investigation of electrical circuits. Therefore, the Laplace transformation plays a key role in electrical engineering.

1.9 Gauge Theory – Local Symmetry and the Description of Interactions by Gauge Fields

As we have discussed in Chap. 2 of Vol. I, the Standard Model in particle physics is based on

- 12 basic particles (6 quarks and 6 leptons), and
- 12 interacting particles (the photon, the 3 vector bosons W^+, W^-, Z^0 and 8 gluons).

This model was formulated in the 1960s and early 1970s. Note the following crucial fact about the structure of the fundamental interactions in nature.

> *The fields of the interacting particles can be obtained from the fields of the basic particles by using the principle of local symmetry (also called the gauge principle).*

Prototype of a gauge theory. Let us explain the basic ideas by considering the following simple model. To this end, let us choose the unit square $Q := \{(x,t) : 0 \le x, t \le 1\}$. We start with the principle of critical action

$$\boxed{\int_Q L(\psi, \psi_t, \psi_x; \psi^\dagger, \psi_t^\dagger, \psi_x^\dagger)\, dxdt = \text{critical!}} \tag{1.35}$$

with the boundary condition $\psi = \psi_0$ on ∂Q and the special Lagrangian

$$L := \psi^\dagger \psi_t + \psi^\dagger \psi_x. \tag{1.36}$$

Here, ψ_t (resp. ψ_x) denotes the partial derivative of ψ with respect to time t (resp. position x). We are given a fixed continuous function $\psi_0 : \partial Q \to \mathbb{C}$ on the boundary of the square Q. We are looking for a smooth function $\psi : Q \to \mathbb{C}$ which solves the variational problem (1.35).

By a basic result from the calculus of variations, we get the following. If the function ψ is a solution of (1.35), then it is a solution of the two Euler–Lagrange equations

$$\frac{\partial}{\partial t} L_{\psi_t^\dagger} + \frac{\partial}{\partial x} L_{\psi_x^\dagger} = L_{\psi^\dagger} \tag{1.37}$$

and

$$\frac{\partial}{\partial t} L_{\psi_t} + \frac{\partial}{\partial x} L_{\psi_x} = L_{\psi}. \tag{1.38}$$

Here, the symbol L_ψ (resp. $L_{\psi^\dagger}$) denotes the partial derivative of L with respect to the variable ψ (resp. $\psi^\dagger$). The proof can be found in Problem 14.7 of Vol. I. Explicitly, the two Euler–Lagrange equations read as

$$\psi_t + \psi_x = 0, \qquad \psi_t^\dagger + \psi_x^\dagger = 0. \tag{1.39}$$

If the function ψ is a solution of (1.39), then we have

$$(\psi\psi^\dagger)_t + (\psi\psi^\dagger)_x = 0, \tag{1.40}$$

which is called a conservation law. In fact, $\psi_t\psi^\dagger + \psi_t^\dagger\psi = -\psi_x\psi^\dagger - \psi_x^\dagger\psi$. This is equal to $-(\psi\psi^\dagger)_x$. Conservation laws play a fundamental role in all fields of physics, since they simplify the computation of solutions. In the 18th and 19th century, astronomers unsuccessfully tried to find $6N$ conservation laws for the motion of N bodies in celestial mechanics ($N \geq 3$), in order to compute the solution and to prove the stability of our solar system.[21]

Step by step, mathematicians and physicists discovered that

Conservation laws are intimately related to symmetries.

The precise formulation of this principle is the content of the Noether theorem proved in 1918 (see Sect. 6.6). We want to show that the invariance of the Lagrangian L (with respect to a global gauge transformation) is behind the conservation law (1.40).

(i) Global symmetry and the Noether theorem. Let α be a fixed real number. We consider the global symmetry transformation

$$\psi_+(x,t) := \mathrm{e}^{\mathrm{i}\alpha}\psi(x,t) \qquad \text{for all} \quad x,t \in \mathbb{R}, \tag{1.41}$$

that is, the field ψ is multiplied by the constant phase factor $\mathrm{e}^{\mathrm{i}\alpha}$, where α is called the phase. The transformation (1.41) is also called a global gauge transformation, by physicists. We also define the infinitesimal gauge transformation $\delta\psi$ by setting

$$\delta\psi(x,t) := \frac{d}{d\alpha}\left(\mathrm{e}^{\mathrm{i}\alpha}\psi(x,t)\right)_{|\alpha=0} = \mathrm{i}\psi(x,t).$$

This means that $\psi_+(x,t) = 1 + \alpha\cdot\delta\psi(x,t) + O(\alpha^2)$ as $\alpha \to 0$. Noting that $\psi_+^\dagger = \mathrm{e}^{-\mathrm{i}\alpha}\psi^\dagger$, the special Lagrangian L from (1.36) is invariant under the global gauge transformation (1.41), that is,

$$\psi_+^\dagger(\psi_+)_t + \psi_+^\dagger(\psi_+)_x = \psi^\dagger\psi_t + \psi^\dagger\psi_x.$$

Generally, the Lagrangian L is invariant under the global gauge transformation (1.41) iff

$$L(\psi_+, (\psi_+)_t, (\psi_+)_x; \psi_+^\dagger, (\psi_+^\dagger)_t, (\psi_+^\dagger)_x) = L(\psi, \psi_t, \psi_x; \psi^\dagger, \psi_t^\dagger, \psi_x^\dagger).$$

[21] See D. Boccaletti and G. Pucacco, Theory of Orbits, Vol 1: Integrable Systems and Non-Perturbative Methods, Vol. 2: Perturbative and Geometrical Methods, Springer, Berlin, 1996.
Y. Hagihara, Celestial Mechanics, Vols. 1–5, MIT Press, Cambridge, Massachusetts, 1976.
W. Neutsch and K. Scherer, Celestial Mechanics: An Introduction to Classical and Contemporary Methods, Wissenschaftsverlag, Mannheim, 1992.

Then a special case of the famous Noether theorem on page 387 tells us the following: If the function ψ is a solution of the variational problem (1.35), then

$$\frac{\partial}{\partial t}\left(L_{\psi_t}\delta\psi + L_{\psi_t^\dagger}\delta\psi^\dagger\right) + \frac{\partial}{\partial x}\left(L_{\psi_x}\delta\psi + L_{\psi_x^\dagger}\delta\psi^\dagger\right) = 0.$$

If we choose the special Lagrangian $L = \psi^\dagger\psi_t + \psi^\dagger\psi_x$, then we obtain the conservation law (1.40).

(ii) Local symmetry and the covariant derivative. We now replace the global gauge transformation (1.40) by the following local gauge transformation

$$\boxed{\psi_+(x,t) := \mathrm{e}^{\mathrm{i}\alpha(x,t)}\psi(x,t) \qquad \text{for all} \quad x,t \in \mathbb{R},} \tag{1.42}$$

where the phase α depends on space and time. We postulate the following crucial local symmetry principle:

(P) *The Lagrangian L is invariant under local gauge transformations.*

It can be easily shown that the function L from (1.36) does *not* possess this invariance property for arbitrary functions $\alpha = \alpha(x,t)$. This follows from

$$(\psi_+)_t = \mathrm{i}\alpha_t\mathrm{e}^{\mathrm{i}\alpha}\psi + \mathrm{e}^{\mathrm{i}\alpha}\psi_t.$$

Here, the appearance of the derivative α_t of the phase function α destroys the invariance property of L.

Our *goal* is to modify the function L in such a way that it is invariant under (1.42). To this end, we introduce the so-called covariant partial derivatives

$$\boxed{\nabla_t := \frac{\partial}{\partial t} + \mathrm{i}U(x,t), \qquad \nabla_x := \frac{\partial}{\partial x} + \mathrm{i}A(x,t),} \tag{1.43}$$

where $U, A : \mathbb{R}^2 \to \mathbb{R}$ are given smooth real-valued functions called gauge fields. The local gauge transformation of U and A is defined by

$$U_+ := U - \alpha_t, \qquad A_+ := A - \alpha_x.$$

Furthermore, we define the following transformation law for the covariant partial derivatives:

$$\nabla_t^+ := \frac{\partial}{\partial t} + \mathrm{i}U_+, \qquad \nabla_x^+ := \frac{\partial}{\partial x} + \mathrm{i}A_+. \tag{1.44}$$

The key relation is given by the following elegant transformation law for the covariant partial derivatives:

$$\boxed{\nabla_t^+\psi_+ = \mathrm{e}^{\mathrm{i}\alpha}\nabla_t\psi, \qquad \nabla_x^+\psi_+ = \mathrm{e}^{\mathrm{i}\alpha}\nabla_x\psi.} \tag{1.45}$$

Theorem 1.1 *There holds* (1.45).

This theorem tells us the crucial fact that, in contrast to the classical partial derivatives, the covariant partial derivatives are transformed in the same way as the field ψ itself. This property is typical for covariant partial derivatives in mathematics. Indeed, our construction of covariant partial derivatives has been chosen in such a way that (1.45) is valid.
Proof. By the product rule,

$$\left(\frac{\partial}{\partial t} + \mathrm{i}U_+\right)\psi_+ = \mathrm{e}^{\mathrm{i}\alpha}(\mathrm{i}\alpha_t\psi + \psi_t + \mathrm{i}U_+\psi) = \mathrm{e}^{\mathrm{i}\alpha}\left(\frac{\partial}{\partial t} + \mathrm{i}U\right)\psi.$$

This yields $\nabla_t^+\psi_+ = \mathrm{e}^{\mathrm{i}\alpha}\nabla_t\psi$. Similarly, we get $\nabla_x^+\psi_+ = \mathrm{e}^{\mathrm{i}\alpha}\nabla_x\psi$. □
Now let us discuss the main idea of gauge theory:

We replace the classical partial derivatives $\frac{\partial}{\partial t}, \frac{\partial}{\partial x}$ *by the covariant partial derivatives* ∇_t, ∇_x*, respectively.*

This is the main trick of gauge theory. In particular, we replace the Lagrangian

$$L = \psi^\dagger\frac{\partial}{\partial t}\psi + \psi^\dagger\frac{\partial}{\partial x}\psi$$

from the original variational problem (1.35) by the modified Lagrangian

$$\boxed{L := \psi^\dagger\nabla_t\psi + \psi^\dagger\nabla_x\psi.}$$

Explicitly, we have

$$L = \psi^\dagger\psi_t + \psi^\dagger\psi_x + \mathrm{i}\psi^\dagger U\psi + \mathrm{i}\psi^\dagger A\psi.$$

The corresponding Euler–Lagrange equations (1.37) and (1.38) read as

$$\nabla_t\psi + \nabla_x\psi = 0, \tag{1.46}$$

and $(\nabla_t\psi + \nabla_x\psi)^\dagger = 0$, respectively.

The local symmetry principle (P) above is closely related to the Faraday–Green locality principle, saying that physical interactions are localized in space and time.

Summarizing, the local symmetry principle (P) *enforces the existence of additional gauge fields* U, A *which interact with the originally given field* ψ*.*

In the Standard Model in particle physics and in the theory of general relativity, the additional gauge fields are responsible for the interacting particles.

Consequently, the mathematical structure of the fundamental interactions in nature is a consequence of the local symmetry principle.

In his search of a unified theory for all interactions in nature, Einstein was not successful, since he was not aware of the importance of the principle of local symmetry. In our discussion below, the following notions will be crucial:

- local gauge transformation,
- gauge force F,
- connection form $\mathcal{A}$,
- curvature form $\mathcal{F}$ (gauge force form), and
- parallel transport of information.

Gauge force. Covariant partial derivatives can be used in order to introduce the following notions:

(a) Gauge force (also called curvature): We define

$$\boxed{\mathrm{i}F := \nabla_x \nabla_t - \nabla_t \nabla_x.} \tag{1.47}$$

In physics, the function F is called the **force** induced by the gauge fields U, A. Explicitly, we get

$$F = U_x - A_t. \tag{1.48}$$

Relation (1.47) tells us that:

The "gauge force" F measures the non-commutativity of the covariant partial derivatives.

In particular, the force F vanishes if the gauge fields U, A vanish. The proof of (1.48) follows from

$$\begin{aligned}\nabla_t(\nabla_x\psi) &= \left(\frac{\partial}{\partial t} + \mathrm{i}U\right)(\psi_x + \mathrm{i}A\psi)\\ &= \psi_{tx} + \mathrm{i}A_t\psi + \mathrm{i}A\psi_t + \mathrm{i}U\psi_x - UA\psi\end{aligned}$$

and

$$\begin{aligned}\nabla_x(\nabla_t\psi) &= \left(\frac{\partial}{\partial x} + \mathrm{i}A\right)(\psi_t + \mathrm{i}U\psi)\\ &= \psi_{xt} + \mathrm{i}U_x\psi + \mathrm{i}U\psi_x + \mathrm{i}A\psi_t - AU\psi.\end{aligned}$$

Hence $(\nabla_x\nabla_t - \nabla_t\nabla_x)\psi = \mathrm{i}(U_x - A_t)\psi$. □

The transformation of the force F with respect to the gauge transformation $\psi_+(x,t) = \mathrm{e}^{\mathrm{i}\alpha(x,y)}\psi(x,t)$ is defined by

$$\mathrm{i}F^+ := \nabla_x^+\nabla_t^+ - \nabla_t^+\nabla_x^+.$$

Theorem 1.2 $F^+ = e^{\mathrm{i}\alpha}Fe^{-\mathrm{i}\alpha}$.

Proof. It follows from Theorem 1.1 on page 37 that

$$\begin{aligned} \mathrm{i}F^+\psi_+ &= (\nabla_x^+\nabla_t^+ - \nabla_t^+\nabla_x^+)\psi_+ = \nabla_x^+(\mathrm{e}^{\mathrm{i}\alpha}\nabla_t\psi) - \nabla_t^+(\mathrm{e}^{\mathrm{i}\alpha}\nabla_x\psi) \\ &= \mathrm{e}^{\mathrm{i}\alpha}(\nabla_x\nabla_t\psi - \nabla_t\nabla_x\psi) = \mathrm{e}^{\mathrm{i}\alpha}\mathrm{i}F\psi = (\mathrm{e}^{\mathrm{i}\alpha}\mathrm{i}F\mathrm{e}^{-\mathrm{i}\alpha})\psi_+. \end{aligned}$$

□

In the present case, we have the commutativity property $F\mathrm{e}^{-\mathrm{i}\alpha} = \mathrm{e}^{-\mathrm{i}\alpha}F$. Hence

$$F^+ = \mathrm{e}^{\mathrm{i}\alpha}\mathrm{e}^{-\mathrm{i}\alpha}F = F,$$

that is, the force F is gauge invariant. In more general gauge theories, the phase factor $\mathrm{e}^{\mathrm{i}\alpha(x,t)}$ is a matrix. In this case, the force F is not gauge invariant anymore. However, it is possible to construct gauge invariants which depend on F. This is the case for the Standard Model in particle physics (see Vol. III).

(b) Covariant directional derivative: Consider the curve

$$C : x = x(\sigma), \quad t = t(\sigma),$$

where the curve parameter σ varies in the interval $[0, \sigma_0]$. The classical directional derivative along the curve C is defined by

$$\frac{d}{d\sigma} := \frac{dx(\sigma)}{d\sigma}\frac{\partial}{\partial x} + \frac{dt(\sigma)}{d\sigma}\frac{\partial}{\partial t}.$$

Explicitly, we get

$$\frac{d}{d\sigma}\psi(x(\sigma), t(\sigma)) = \frac{dx(\sigma)}{d\sigma}\psi_x(x(\sigma), t(\sigma)) + \frac{dt(\sigma)}{d\sigma}\psi_t(x(\sigma), t(\sigma)).$$

Similarly, the covariant directional derivative along the curve C is defined by

$$\boxed{\frac{D}{d\sigma} := \frac{dx(\sigma)}{d\sigma}\nabla_x + \frac{dt(\sigma)}{d\sigma}\nabla_t.}$$

Explicitly,

$$\begin{aligned} \frac{D}{d\sigma}\psi(x(\sigma), t(\sigma)) = \frac{d}{d\sigma}\psi(x(\sigma), t(\sigma)) + \mathrm{i}A(x(\sigma)), t(\sigma))\frac{dx(\sigma)}{d\sigma} \\ +\mathrm{i}U(x(\sigma), t(\sigma))\frac{dt(\sigma)}{d\sigma}. \end{aligned} \tag{1.49}$$

(c) Parallel transport: We say that the field function ψ is parallel along the curve C iff

$$\boxed{\frac{D}{d\sigma}\psi(x(\sigma), t(\sigma)) = 0, \qquad 0 \le \sigma \le \sigma_0.} \tag{1.50}$$

By (1.49), this notion depends on the gauge fields U, A. In particular, if the gauge fields U, A vanish, then parallel transport means that the field ψ is constant along the curve C.

The following observation is crucial. It follows from the key relation (1.45) on page 36 that the equation (1.50) of parallel transport is invariant under local gauge transformations. This means that (1.50) implies

$$\frac{D^+}{d\sigma}\,\psi_+(x(\sigma),t(\sigma)) = 0, \qquad 0 \le \sigma \le \sigma_0.$$

Consequently, in terms of mathematics, parallel transport possesses a geometric meaning with respect to local symmetry transformations.

In terms of physics, parallel transport describes the transport of physical information in space and time.

This transport is local in space and time, which reflects the Faraday–Green locality principle.

The Cartan differential. The most elegant formulation of gauge theories is based on the use of the covariant Cartan differential. As a preparation, let us recall the classical Cartan calculus. We will use the following relations:

$$dx \wedge dt = -dt \wedge dx, \quad dx \wedge dx = 0, \quad dt \wedge dt = 0. \tag{1.51}$$

Moreover, the wedge product of three factors of the form dx, dt is always equal to zero. For example,

$$dx \wedge dt \wedge dt = 0, \quad dt \wedge dx \wedge dt = 0. \tag{1.52}$$

For the wedge product, both the distributive law and the associative law are valid. Let $\psi : \mathbb{R}^2 \to \mathbb{C}$ be a smooth function. By definition,

- $d\psi := \psi_x dx + \psi_t dt$.

The differential 1-form

$$\boxed{\mathcal{A} := \mathrm{i}A dx + \mathrm{i}U dt} \tag{1.53}$$

is called the Cartan **connection form**. By definition,

- $d\mathcal{A} := \mathrm{i}dA \wedge dx + \mathrm{i}dU \wedge dt$,
- $d(\psi\, dx \wedge dt) = d\psi \wedge dx \wedge dt = 0$ (Poincaré identity).

The Poincaré identity is a consequence of (1.52).

The covariant Cartan differential. We now replace the classical partial derivatives by the corresponding covariant partial derivatives. Therefore, we replace $d\psi$ by the definition

- $D\psi := \nabla_x \psi\, dx + \nabla_t \psi\, dt$.

Similarly, we define

- $D\mathcal{A} := \mathrm{i}DA \wedge dx + \mathrm{i}DU \wedge dt$,
- $D(\psi\, dx \wedge dt) = D\psi \wedge dx \wedge dt = 0$ (Bianchi identity).

The Bianchi identity is a consequence of (1.52). Let us introduce the Cartan **curvature form** $\mathcal{F}$ by setting

$$\boxed{\mathcal{F} := D\mathcal{A}.} \tag{1.54}$$

Theorem 1.3 (i) $D\psi = d\psi + \mathcal{A}\psi$.
(ii) $\mathcal{F} = d\mathcal{A} + \mathcal{A} \wedge \mathcal{A}$ *(Cartan's structural equation).*
(iii) $D\mathcal{F} = 0$ *(Bianchi identity).*

In addition, we have the following relations for the curvature form $\mathcal{F}$:

- $\mathcal{F} = d\mathcal{A} + [U, A]_-$.[22]
- $\mathcal{F} = \mathrm{i}F dx \wedge dt$, where $\mathrm{i}F = U_x - A_t$.

Proof. Ad (i). $D\psi = (\psi_x + \mathrm{i}A\psi)dx + (\psi_t + \mathrm{i}U\psi)dt$.
Ad (ii). Note that

$$\begin{aligned} DA &= \mathrm{i}(A_x + A^2)dx + \mathrm{i}(A_t + \mathrm{i}UA)dt, \\ DU &= \mathrm{i}(U_x + \mathrm{i}AU)dx + (U_t + \mathrm{i}U^2)dt, \end{aligned}$$

and

$$\mathcal{A} \wedge \mathcal{A} = -(Adx + Udt) \wedge (Adx + Udt) = (UA - AU)\, dx \wedge dt.$$

Hence

$$\begin{aligned} D\mathcal{A} &= \mathrm{i}DA \wedge dx + \mathrm{i}DU \wedge dt \\ &= \mathrm{i}(A_t + \mathrm{i}UA)\, dt \wedge dx + \mathrm{i}(U_x + \mathrm{i}AU)\, dx \wedge dt = d\mathcal{A} + \mathcal{A} \wedge \mathcal{A}. \end{aligned}$$

This yields all the identities claimed above. □

The results concerning the curvature form $\mathcal{F}$ above show that Cartan's structural equation (1.54) is nothing else than a reformulation of the equation

$$\mathrm{i}F = U_x - A_t,$$

which relates the force F to the potentials U, A. Furthermore, we will show in Sect. 5.11 on page 333 that Cartan's structural equation is closely related to both

- Gauss' theorema egregium on the computation of the Gaussian curvature of a classic surfaces by means of the metric tensor and its partial derivatives,
- and the Riemann formula for the computation of the Riemann curvature tensor of a Riemannian manifold by means of the metric tensor and its partial derivatives.

In the present case, the formulas can be simplified in the following way. It follows from the commutativity property $AU = UA$ that:

[22] Here, we use the Lie bracket $[U, A]_- := UA - AU$.

- $\mathcal{F} = dA$,
- $\mathcal{F} = \mathrm{i}F\ dx \wedge dt = \mathrm{i}(U_x - A_t)\ dx \wedge dt$.

A similar situation appears in Maxwell's theory of electromagnetism. For more general gauge theories, the symbols A and U represent matrices. Then we obtain the additional nonzero terms $[U, A]_-$ and $\mathcal{A} \wedge \mathcal{A}$. This is the case in the Standard Model of elementary particles (see Vol. III).

The mathematical language of fiber bundles. In mathematics, we proceed as follows:

- We consider the field $\psi : \mathbb{R}^2 \to \mathbb{C}$ as a section of the line bundle $\mathbb{R}^2 \times \mathbb{C}$ (with typical fiber $\mathbb{C}$) (see Fig. 4.9 on page 208).
- The line bundle $\mathbb{R}^2 \times \mathbb{C}$ is associated to the principal fiber bundle $\mathbb{R}^2 \times U(1)$ (with structure group $U(1)$ called the gauge group in physics).[23]
- As above, the differential 1-form $\mathcal{A} := \mathrm{i}A dx + \mathrm{i}U dt$ is called the connection form on the base manifold $\mathbb{R}^2$ of the principal fiber bundle $\mathbb{R}^2 \times U(1)$, and
- the differential 2-form

$$\mathcal{F} = d\mathcal{A} + \mathcal{A} \wedge \mathcal{A}$$

  is called the curvature form on the base manifold $\mathbb{R}^2$ of the principle fiber bundle $\mathbb{R}^2 \times U(1)$.
- Finally, we define

$$D\psi := d\psi + \mathcal{A}\psi. \tag{1.55}$$

  This is called the covariant differential of the section ψ of the line bundle $\mathbb{R}^2 \times \mathbb{C}$.

Observe that:

> *The values of the gauge field functions* $\mathrm{i}U, \mathrm{i}A$ *are contained in the Lie algebra* $u(1)$ *of the Lie group* $U(1)$*. Thus, the connection form* $\mathcal{A}$ *is a differential 1-form with values in the Lie algebra* $u(1)$*.*

This can be generalized by replacing

- the special commutative Lie group $U(1)$
- by the the general Lie group $\mathcal{G}$.

Then the values of the gauge fields $\mathrm{i}U, \mathrm{i}A$ are contained in the Lie algebra $\mathcal{LG}$ to $\mathcal{G}$. If $\mathcal{G}$ is a noncommutative Lie group (e.g., $SU(N)$ with $N \geq 2$), then the additional force term $\mathcal{A} \wedge \mathcal{A}$ does not vanish identically, as in the special case of the commutative group $U(1)$.[24] In Vol. III on gauge theory, we will show that the Standard Model in particle physics corresponds to this approach by choosing the gauge group $U(1) \times SU(2) \times SU(3)$. Here,

[23] Recall that the elements of the Lie group $U(1)$ are the complex numbers $\mathrm{e}^{\mathrm{i}\alpha}$ with real parameter α. The elements of the Lie algebra $u(1)$ are the purely imaginary numbers $\alpha\mathrm{i}$.

[24] The Lie group $SU(N)$ consists of all the unitary $(N \times N)$-matrices whose determinant is equal to one (special unitary group).

- the electroweak interaction is the curvature of a $(U(1) \times SU(2))$-bundle (Glashow, Salam and Weinberg in the 1960s), and
- the strong interaction is the curvature of a $SU(3)$-bundle (Gell-Mann and Fritzsch in the early 1970s).

Historical remarks. General gauge theory is equivalent to modern differential geometry. This will be thoroughly studied in Vol. III. At this point let us only make a few historical remarks.

> *In 1827 Gauss proved that the curvature of a 2-dimensional surface in 3-dimensional Euclidean space is an intrinsic property of the manifold.*

This means that the curvature of the surface can be measured without using the surrounding space. This is the content of Gauss' *theorema egregium*. The Gauss theory was generalized to higher-dimensional manifolds by Riemann in 1854. Here, the Gaussian curvature has to be replaced by the Riemann curvature tensor. In 1915 Einstein used this mathematical approach in order to formulate his theory of gravitation (general theory of relativity). In Einstein's setting, the masses of celestial bodies (stars, planets, and so on) determine the Riemann curvature tensor of the four-dimensional space-time manifold which corresponds to the universe. Thus, Newton's gravitational force is replaced by the curvature of a four-dimensional pseudo-Riemannian manifold $\mathcal{M}^4$. The motion of a celestial body (e.g., the motion of a planet around the sun) is described by a geodesic curve C in $\mathcal{M}^4$. Therefore, Einstein's equation of motion tells us that the 4-dimensional velocity vector of C is parallel along the curve C. Roughly speaking, this corresponds to (1.50) where ψ has to be replaced by the velocity field of C. In the framework of his theory of general relativity, Einstein established the principle

$$\textit{force} = \textit{curvature}$$

for gravitation. Nowadays, the Standard Model in particle physics is also based on this beautiful principle which is the most profound connection between mathematics and physics.

In 1917 Levi-Civita introduced the notion of parallel transport, and he showed that both the Gaussian curvature of 2-dimensional surfaces and the Riemann curvature tensor of higher-dimensional manifolds can be computed by using parallel transport of vector fields along small closed curves. In the 1920s, Élie Cartan invented the method of moving frames.[25] In the 1950s, Ehresmann generalized Cartan's method of moving frames to the modern curvature theory for principal fiber bundles (i.e., the fibers are Lie groups) and their associated vector bundles (i.e., the fibers are linear spaces). In 1963, Kobayashi and Nomizu published the classic monograph

[25] For an introduction to this basic tool in modern differential geometry, we refer to the textbook by T. Ivey and J. Landsberg, Cartan for Beginners: Differential Geometry via Moving Frames and Exterior Differential Systems, Amer. Math. Soc., Providence, Rhode Island, 2003. See also Vol. III.

Foundations of Differential Geometry,

Vols. 1, 2, Wiley, New York. This finishes a longterm development in mathematics.

In 1954, the physicists Yang and Mills created the Yang–Mills theory. It was their goal to generalize Maxwell's electrodynamics. To this end, they started with the observation that Maxwell's electrodynamics can be formulated as a gauge theory with the gauge group $U(1)$. This was known from Hermann Weyl's paper: Elektron und Gravitation, Z. Phys. **56** (1929), 330–352 (in German). Yang and Mills

- replaced the commutative group $U(1)$
- by the non-commutative group $SU(2)$.

The group $SU(2)$ consists of all the complex (2×2)-matrices A with $AA^\dagger = I$ and $\det A = 1$. Interestingly enough, in 1954 Yang and Mills did not know a striking physical application of their model. However, in the 1960s and 1970s, the Standard Model in particle physics was established as a modified Yang–Mills theory with the gauge group

$$\boxed{U(1) \times SU(2) \times SU(3).}$$

The modification concerns the use of an additional field called Higgs field in order to generate the masses of the three gauge bosons W^+, W^-, Z^0. In the early 1970s, Yang noticed that the Yang–Mills theory is a special case of Ehresmann's modern differential geometry in mathematics. For the history of gauge theory, we refer to:

L. Brown et al. (Eds.), The Rise of the Standard Model, Cambridge University Press, 1995.

L. O'Raifeartaigh, The Dawning of Gauge Theory, Princeton University Press, 1997.

C. Taylor (Ed.), Gauge Theories in the Twentieth Century, World Scientific, Singapore, 2001 (a collection of fundamental articles).

Mathematics and physics. Arthur Jaffe writes the following in his beautiful survey article *Ordering the universe: the role of mathematics* in the Notices of the American Mathematical Society **236** (1984), 589–608:[26]

> There is an exciting development taking place right now, *reunification* of mathematics with theoretical physics... In the last ten or fifteen years mathematicians and physicists realized that modern geometry is in fact the natural framework for gauge theory. The gauge potential in gauge theory is the connection of mathematics. The gauge field is the mathematical curvature defined by the connection; certain *charges* in physics are the topological invariants studied by mathematicians. While the mathematicians and physicists worked separately on similar ideas, they did *not*

[26] Reprinted by permission of the American Mathematical Society. This report was originated by the National Academy of Sciences of the U.S.A.

> duplicate each other's efforts. The mathematicians produced general, far-reaching theories and investigated their ramifications. Physicists worked out details of certain examples which turned out to describe nature beautifully and elegantly. When the two met again, the results are more powerful than either anticipated... In mathematics, we now have a new motivation to use specific insights from the examples worked out by physicists. This signals the return to an ancient tradition.

Felix Klein (1849–1925) writes about mathematics:

> Our science, in contrast to others, is not founded on a single period of human history, but has accompanied the development of culture through all its stages. Mathematics is as much interwoven with Greek culture as with the most modern problems in engineering. It not only lends a hand to the progressive natural sciences but participates at the same time in the abstract investigations of logicians and philosophers.

Hints for further reading:

S. Chandrasekhar, Truth and Beauty: Aesthetics and Motivations in Science, Chicago University Press, Chicago, Illinois, 1990.

E. Wigner, Philosophical Reflections and Syntheses. Annotated by G. Emch. Edited by J. Mehra and A. Wightman, Springer, New York, 1995.

G. 't Hooft, In Search for the Ultimate Building Blocks, Cambridge University Press, 1996.

R. Brennan, Heisenberg Probably Slept Here: The Lives, Times, and Ideas of the Great Physicists of the 20th Century, Wiley, New York, 1997.

J. Wheeler and K. Ford, Geons, Black Holes, and Quantum Foam: a Life in Physics, Norton, New York, 1998.

B. Greene, The Elegant Universe: Supersymmetric Strings, Hidden Dimensions, and the Quest for the Ultimate Theory, Norton, New York, 1999.

A. Zee, Fearful Symmetry: The Search for Beauty in Modern Physics, Princeton University Press, 1999.

G. Johnson, Strange Beauty: Murray Gell-Mann and the Revolution in Twentieth Century Physics, A. Knopf, New York, 2000.

G. Farmelo (Ed.), It Must be Beautiful: Great Equations of Modern Science, Granta Publications, London, 2003.

M. Veltman, Facts and Mysteries in Elementary Particle Physics, World Scientific, Singapore, 2003.

R. Penrose, The Road to Reality: A Complete Guide to the Laws of the Universe, Jonathan Cape, London, 2004.

J. Barrow, New Theories of Everything: The Quest for Ultimate Explanation, Oxford University Press, New York, 2007.

F. Patras, La pensée mathématique contemporaine (Philosophy of modern mathematics), Presses Universitaire de France, Paris, 2001 (in French).

H. Wußing, 6000 Years of Mathematics: a Cultural Journey through Time, Vols. I, II, , Springer, Heidelberg, 2008 (in German).

E. Zeidler, Reflections on the future of mathematics. In: H. Wußing (2008), Vol. II (last chapter) (in German).

The Cambridge Dictionary of Philosophy, edited by R. Audi, Cambridge University Press, 2005.

1.10 The Challenge of Dark Matter

> Although science teachers often tell their students that the periodic table of the elements shows what the Universe is made of, this is not true. We now know that most of the universe – about 96% of it – is made of dark matter that defies brief description, and certainly is not represented by Mendeleev's periodic table. This unseen 'dark matter' is the subject of this book...
> Dark matter provides a further remainder that we humans are not essential to the Universe. Ever since Copernicus (1473–1543) and others suggested that the Earth was not the center of the Universe, humans have been on a slide away from cosmic significance. At first we were not at the center of the Solar System, and then the Sun became just another star in the Milky Way, not even in the center of our host Galaxy. By this stage the Earth and its inhabitants had vanished like a speck of dust in a storm. This was a shock.
> In the 1930s Edwin Hubble showed that the Milky Way, vast as it is, is a mere 'island Universe' far removed from everywhere special; and even our home galaxy was suddenly insignificant in a sea of galaxies, then clusters of galaxies. Now astronomers have revealed that we are not even made of the same stuff as most of the Universe. While our planet – our bodies, even – are tangible and visible, most of the matter in the Universe is not. Our Universe is made of darkness. How do we respond to that?
>
> Ken Freeman and Geoff McNamarra, 2006

This quotation is taken from the monograph by K. Freemann and G. McNamarra, In Search of Dark Matter, Springer, Berlin and Praxis Publishing Chichester, United Kingdom, 2006 (reprinted with permission). As an introduction to modern cosmology we recommend the monograph by S. Weinberg, Cosmology, Oxford University, 2008.