

Prologue

Geometry is the knowledge of what eternally exists.

Plato of Athen (428–348 B.C.)

He who understands geometry may understand anything in this world.

Galileo Galilei (1564–1642)

The way of people to the laws of nature are not less admirable than the laws themselves.

Johannes Kepler (1571–1630)

In humbleness, we have to admit that if ‘number’ is a product of our imagination, ‘space’ has a reality outside of our imagination, to which *a priori* we cannot assign its laws.

Gauss (1777–1855) in a letter to Bessel, 1840

This prologue should help the reader to understand the sophisticated historical development of gauge theory in mathematics and physics. We will not follow a strict logical route. This will be done later on. At this point, we are going to emphasize the basic ideas. It is our goal to show the reader how the methods of modern differential geometry work in the case of Einstein’s theory of general relativity, which describes the gravitational force in nature. In particular, we want to show how

- the language of physicists created by Einstein and used in most physics textbooks (based on the use of local space-time coordinates) and
- the language of mathematicians used in modern textbooks on differential geometry (based on the invariant – i.e., coordinate-free – formulation)

are related to each other. This should help physicists to enter modern differential geometry. One cannot grasp modern physics without understanding gauge field theory which tells us the following crucial facts:

- interactions in nature are based on the *parallel transport* of physical information;
- forces are described by *curvature* which measures the path-dependence of the parallel transport.

Here, we will discuss the following points:

- an interview with the Nobel prize laureate Chen Ning Yang (born 1922) on the history of modern gauge theory,
- Einstein’s theory of general relativity on gravitation,
- changing observers in the universe and tensor calculus,
- the Riemann curvature tensor and the beauty of Gauss’ theorema egregium,
- two fundamental variational principles in general relativity,

- symmetry and Felix Klein's invariance principle in geometry (a glance at the history of invariant theory in the 19th century),¹
- Einstein's principle of general relativity and invariants – the geometrization of physics (the paradigm of higher-dimensional cartography),
- gauge transformations:
 - Einstein's gauge transformation in the theory of both special relativity and general relativity (change of the observer),
 - Dirac's unitary gauge transformations in the Hilbert space approach to quantum mechanics (change of the observer by changing the measurement device),
 - Yang's gauge transformation by changing the local phase factor of the wave function,
 - the $U(1)$ -gauge transformation in classical electrodynamics and quantum electrodynamics,
 - the $U(1) \times SU(2)$ gauge transformations in electroweak interaction,
 - the $SU(3)$ gauge transformations in strong interaction (quantum chromodynamics),
 - the $U(1) \times SU(2) \times SU(3)$ gauge transformations in the Standard Model in particle physics,
 - the conformal gauge transformations in string theory,
 - Élie Cartan's gauge transformations in his method of moving frames (change of the frame),
- construction of invariants by the universal index killing principle,
- Lie's intrinsic tangent vectors,
- Élie Cartan's algebraization of calculus and infinitesimals,
- Riemann's invariant sectional curvature and the geometric meaning of Riemann's curvature tensor,
- Levi-Civita's parallel transport and the geometric meaning of the Riemann curvature tensor,
- two fundamental approaches in differential geometry:
 - Gauss' method of symmetric tensors, and
 - Cartan's method of antisymmetric tensors,
- Yang's matrix trick (the relation between the Einstein equations in general relativity and the Maxwell–Yang–Mills equations), and Cartan's calculus for matrices with differential forms as entries,
- Cartan's structural equations:
 - local structural equations,
 - global structural equations,
- partial covariant derivative and the classical Ricci calculus,
- the Lie structure behind curvature,
- the generalized Riemann curvature tensor in modern mathematics and physics,
- parallel transport of physical information and curvature,
- the modern language of fiber bundles in mathematics and physics,
- summary of typical applications,
- perspectives (instantons and gauge theory, conformal symmetry and twistors, the Seiberg–Witten equations and the quark confinement, the Donaldson theory for 4-dimensional manifolds, Morse theory and Floer homology, quantum cohomology, J -holomorphic curves, Frobenius manifolds, Ricci flow and the Poincaré conjecture).

¹ One has to distinguish between the German mathematician Felix Klein (1849–1925) and the Swedish physicist Oskar Klein (1894–1977) (one of the authors of the Klein–Fock–Gordon equation in quantum field theory).

The classical formulas (0.13) and (0.14) on page 11 for defining the Riemann curvature tensor via Christoffel symbols for the metric tensor are clumsy. The development of modern differential geometry was essentially influenced by the desire of mathematicians to get insight into the true structure of curvature. This led to a better understanding of curvature and to far-reaching generalizations which proved to be useful in modern physics. The basic paper in mathematics is due to:

C. Ehresmann, *Les connexions infinitésimales dans un espace fibré différentiable* (in French) (The infinitesimal connections in a differentiable fiber bundle), Colloque de Topologie, Bruxelles, 1950, pp. 29–55.

Charles Ehresmann (1905–1979) based his theory on Élie Cartan’s work created in the 1920s.² The first textbook on modern differential geometry was written by:

S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Vols. 1, 2, Wiley, New York, 1963.

We also recommend:

Y. Choquet-Bruhat, C. DeWitt-Morette, and M. Dillard-Bleick, *Analysis, Manifolds, and Physics*. Vol. 1: Basics; Vol. 2: 92 Applications, Elsevier, Amsterdam, 1996.

T. Frankel, *The Geometry of Physics*, Cambridge University Press, Cambridge, 2004.

S. Novikov and T. Taimanov, *Geometric Structures and Fields*, Amer. Math. Soc., Providence, Rhode Island, 2006.

As an introduction to the theory of general relativity based on the use of local coordinates, we recommend the classical Lecture Notes by

P. Dirac, *General Theory of Relativity*, Princeton University Press, 1996 (70 pages)

together with

Ø. Grøn and S. Hervik, *Einstein’s Theory of General Relativity: with Modern Applications in Cosmology*, Springer, New York, 2007.

Both the invariant formulation and the formulation in terms of local coordinates is discussed in great detail in the classic textbook by

C. Misner, K. Thorne, and A. Wheeler, *Gravitation*, Freeman, San Francisco, California, 1973.

For the sophisticated mathematical problem of solving the initial-value problem for the Einstein equations on the gravitational field, we recommend:

P. Cruściel and H. Friedrich, *The Einstein Equations and the Large Scale Behavior of Gravitational Fields: 50 Years of the Cauchy Problem in General Relativity*, Birkhäuser, Boston, 2004.

Y. Choquet-Bruhat, *General Relativity and the Einstein Equations*, Oxford University Press, 2008.

As a comprehensive modern textbook, we recommend:

² One has to distinguish between the great French geometer Élie Cartan (1869–1951), who strongly influenced the development of modern differential geometry, and his famous son Henri Cartan (1904–2008), who made important contributions to algebra, analysis, and topology (e.g., homological algebra, the theory of analytic functions of several complex variables, and the cohomology of sheaves).

T. Padmanabhan, *Gravitation: Foundations and Frontiers*, Cambridge University Press, 2010.

The nature of dark matter is one of the great open problems in physics. We refer to:

G. Bertone (Ed.), *Particle Dark Matter*, Cambridge University Press, 2010.

An Interview with Chen Ning Yang on the History of Modern Gauge Theory

To begin with, let us quote some parts of an interview given by the physicist Chen Ning Yang answering the questions of Dianzhou Zhang:³

Zhang: Chen Ning Yang (born 1922 in Hefei, China), one of the twentieth century's great theoretical physicists, shared the Nobel prize in physics with Tsung-Dao Lee in 1957 for their joint contribution to parity non-conservation in weak interaction. Mathematicians, however, know Yang best for the Yang–Mills gauge field theory and the Yang–Baxter equation. After Einstein and Dirac, Yang is perhaps the twentieth-century physicist who has had the greatest impact on the development of mathematics ... While a student in Kunming (China) and Chicago, Yang was impressed with the fact that gauge invariance determined all electromagnetic interactions. This was known from the works in the years 1918–1929 of Weyl, Fock, and London, and through later review papers by Pauli. But by the 1940s and the early 1950s, it played only a minor and technical role in physics. In Chicago, Yang tried to generalize the concept of gauge invariance to non-Abelian groups (the gauge group for electromagnetism being the Abelian group $U(1)$). In analogy with Maxwell's equations he tried

$$\mathcal{F}_{\alpha\beta} = \partial_\alpha \mathcal{A}_\beta - \partial_\beta \mathcal{A}_\alpha,$$

where A_α are matrices ($\alpha, \beta = 0, 1, 2, 3$). As Yang pointed out later on, “This led to a mess, and I had to give up.”

In 1954, as a visiting physicist at Brookhaven National Laboratory on Long Island, New York, Yang returned once again to the idea of generalizing gauge invariance. His officemate was Robert Mills, who was about to finish his Ph.D. degree at Columbia University, New York City. Yang introduced the idea of non-Abelian gauge field to Mills, and they decided to add a quadratic term:⁴

$$\boxed{\mathcal{F}_{\alpha\beta} = \partial_\alpha \mathcal{A}_\beta - \partial_\beta \mathcal{A}_\alpha + \mathcal{A}_\alpha \mathcal{A}_\beta - \mathcal{A}_\beta \mathcal{A}_\alpha.} \quad (0.1)$$

That cleared up the “mess” and led to a beautiful new field theory.⁵

Zhang: *Did you study gauge field theory continuously after 1954?*

³ D. Zhang, N. C. Yang and contemporary mathematics, *Mathematical Intelligencer* **15** (1993), Springer, New York, pp. 13–21. Reprinted with permission.

⁴ Here, the point x of the space-time manifold has the (local) coordinates x^0, x^1, x^2, x^3 with $x^0 = ct$ (t time, c velocity of light in a vacuum). Furthermore, the symbol ∂_α denotes the partial derivative $\frac{\partial}{\partial x^\alpha}$.

⁵ C. Yang and R. Mills, Conservation of isotopic spin and isotopic spin invariance, *Phys. Rev.* **96** (1954), 191–195.

Yang: Yes, I did ... In the late 1960s, I began a new formulation of gauge field theory through the approach of non-integrable phase factors. It happened that one semester I was teaching general relativity, and I noticed that the formula (0.1) in gauge field theory and the formula

$$R_{\alpha\beta\gamma}^{\delta} = \partial_{\alpha}\Gamma_{\beta\gamma}^{\delta} - \partial_{\beta}\Gamma_{\alpha\gamma}^{\delta} + \Gamma_{\alpha\mu}^{\delta}\Gamma_{\beta\gamma}^{\mu} - \Gamma_{\beta\mu}^{\delta}\Gamma_{\alpha\gamma}^{\mu} \quad (0.2)$$

with $\alpha, \beta, \gamma, \delta = 0, 1, 2, 3$ for the Riemann curvature tensor in Riemannian geometry are not just similar – they are, in fact, the same if one makes the right identification of symbols.⁶ It is hard to describe the thrill I felt at understanding the point.

Zhang: *Is that the first time that you realized the relation between gauge theory and differential geometry?*

Yang: I had noticed the similarity between Levi-Civita's parallel displacement and non-integrable phase factors in gauge fields. But the exact relationship was appreciated by me only when I realized that the formula (0.1) in gauge field theory and the Riemann formula (0.2) are the same. With an appreciation of the geometrical meaning of gauge theory, I consulted Jim Simons, a distinguished geometer, who was then the chairman of the Mathematics Department at Stony Brooke (Long Island, New York). He said gauge theory must be related to connections on fiber bundles. I then tried to understand fiber-bundle theory from such books as Steenrod's "The Topology of Fiber Bundles," Princeton University Press, 1951, but I learned nothing. The language of modern mathematics is too cold and abstract for a physicist.

Zhang: *I suppose only mathematicians appreciate the mathematical language of today.*

Yang: I can tell you a relevant story. About ten years ago, I gave a talk on physics in Seoul, South Korea. I joked "There exist only two kinds of modern mathematics books: one which you cannot read beyond the first page and one which you cannot read beyond the first sentence. The *Mathematical Intelligencer* later reprinted this joke of mine. But I suspect many mathematicians themselves agree with me.

Zhang: *When did you understand bundle theory?*

Yang: In early 1975, I invited Jim Simons to give us a series of luncheon lectures on differential forms and bundle theory. He kindly accepted the invitation, and we learned about de Rham's theorem, differential forms, patching and so on ...

Zhang: Simon's lecture helped Wu and Yang to write a famous paper in 1975.⁷ In this paper, they analyzed the intrinsic meaning of electromagnetism, emphasizing especially its global topological aspects. They discussed the mathematical meaning of the Aharonov–Bohm experiment and of the Dirac magnetic monopole. They exhibited a dictionary on the translation of terminologies used in mathematics and physics. Half a year later, Isadore Singer of the Massachusetts Institute of Technology (MIT, Cambridge, Massachusetts) visited Stony Brooke and discussed these matters with Yang at length. Singer had been an undergraduate student in physics and a graduate student in mathematics in the 1940s. He wrote in 1985: "Thirty years later I found myself lecturing on gauge theories, beginning with the Wu and Yang dictionary and ending with instantons,

⁶ This will be shown on page 35 under the heading "Yang's matrix trick."

⁷ T. Wu and C. Yang, Concept of non-integrable phase factors and global formulation of gauge fields, *Phys. Rev. D* **12** (1975), 3845–3857.

that is, self-dual connections. I would be inaccurate to say after studying mathematics for thirty years, I felt prepared to return to physics.”

Yang: In 1975, impressed with the fact that gauge fields are connections on fiber bundles, I drove to the house of Shing-Shen Chern (1911–2004) in El Cerrito near Berkeley (California) ... I said I found it amazing that gauge theory are exactly connections on fiber bundles, which the mathematicians developed without reference to the physical world. I added “This is both thrilling and puzzling, since you mathematicians dreamed up these concepts out of nowhere.” Chern immediately protested “No, no. These concepts were not dreamed up. They were natural and real.”

Zhang: *The Yang–Baxter equation*

$$A(u)B(u+v)A(v) = B(v)A(u+v)B(u)$$

appearing in statistical mechanics is just a simple equation for matrix functions. Why does it have such great importance?

Yang: In the simplest situation, the Yang–Baxter equation has the form

$$ABA = BAB.$$

This is the fundamental equation of Artin (1898–1962) for the braid group. The braid group is, of course, a record of the history of permutations. It is not difficult to understand that the history of permutations is relevant to many problems in mathematics and physics. Looking at the developments of the last six or seven years, I got the feeling that the Yang–Baxter equation is the next pervasive algebraic equation after the Jacobi identity

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0.$$

The study of the Jacobi identity has, of course, led to the whole of Lie algebra and its relationship to Lie groups that govern symmetry in nature.

Zhang: Yang–Mills theory and the Yang–Baxter equation both figure prominently in today’s score mathematics. One can see this by the Fields medals awarded in 1986 and 1990. Simon Donaldson was awarded a Fields medal at the International Congress of Mathematicians held in Berkeley in 1986. Sir Michael Atiyah spoke on Simon Donaldson’s work: “Together with the important work of Michael Freedman (another Fields medal winner in 1986), Donaldson’s result implied that there exist ‘exotic’ four-dimensional spaces which are topologically but not differentially equivalent to the standard Euclidean four-dimensional space \mathbb{R}^4 ... Donaldson’s results are derived from the Yang–Mills equations of theoretical physics which are nonlinear generalizations of Maxwell’s equations. In the Euclidean case the solution to the Yang–Mills equations giving the absolute minimum are of special interest and called instantons.”

There were four Fields medalists in 1990: Vladimir Drinfeld, Vaughan Jones, Shigefumi Mori, and Edward Witten. The work of three of them was related to the Yang–Mills equations

$$\boxed{-D * \mathcal{F} = *\mathcal{J}, \quad D\mathcal{F} = 0}$$

and/or the Yang–Baxter equation (see Sect. 15.4).

- (i) We should mention Drinfeld's pioneering work with Yuri Manin on the construction of instantons. These are solutions to the Yang–Mills equations which can be thought of as having particle-like properties of localization and size. Drinfeld's interest in physics continued with his investigation of the Yang–Baxter equation.
- (ii) Jones opened a whole new direction upon realizing that under certain conditions solutions of the Yang–Baxter equation could be used for constructing invariants of links ... The theory of quantum groups (i.e., deformations of classical Lie groups based on non-commutative Hopf algebras) was devised by Jimbo and Drinfeld to produce solutions of Yang–Baxter equations.
- (iii) Witten described in these terms the invariants of Donaldson and Floer (extending the earlier ideas of Atiyah) and generalized the Jones polynomials to the case of an arbitrary ambient three-dimensional manifold.

We note with amusement that there were complaints that the plenary lectures at the International Congress of Mathematicians in Kyoto, 1990, were heavily slanted toward the topics of mathematical physics: "Everywhere we heard quantum group, quantum group, quantum group!" ...

Yang: Many theoretical physicists are, in some ways, antagonistic to mathematics, or at least have a tendency to downplay the value of mathematics. I do not agree with these attitudes. I have written:⁸ "Perhaps of my father's influence, I appreciate mathematics. I appreciate the value judgement of the mathematician, and I admire the beauty and power of mathematics: there are ingenuity and intricacy in tactical maneuvers, and breathtaking sweeps in strategic campaigns. And, of course, miracle of miracles, some concepts in mathematics turn out to provide the fundamental structures that govern the physical universe!"

In the present volume, we will show that the Yang–Mills equations generalize the Maxwell equations in electromagnetism.

Einstein's Theory of General Relativity on Gravitation

We set

$$R_{\alpha\beta} = \kappa_G(T_{\alpha\beta} - \frac{1}{2}g_{\alpha\beta}T), \quad \alpha, \beta = 0, 1, 2, 3.$$

This completes the general theory of relativity as a logical structure. The postulate of relativity in its most general form, which makes the space-time coordinates meaningless parameters, leads necessarily to a certain form of gravitational theory which explains the motion of the Perihelion of the planet Mercury.

Anyone who has really grasped the general theory of relativity, will be captured by its beauty. It is a triumph of the general differential calculus, which was created by Gauss (1777–1855), Riemann (1826–1866), Christoffel (1829–1900), Ricci-Curbastro (1853–1925), Bianchi (1856–1928), and Levi-Civita (1873–1941).⁹

Albert Einstein, 1915

⁸ C. Yang, *Selected Papers*, Freeman, San Francisco, 1983.

⁹ A. Einstein, On general relativity. The field equations of gravitation. Reports on the meetings of the Prussian Academy of Sciences (Berlin) on November 11 and December 2, 1915 (in German).

The two fundamental Einstein equations. In 1915, motivated by the study of classical differential geometry, Einstein based his theory of general relativity on the Riemann curvature tensor of the four-dimensional space-time manifold \mathcal{M}^4 . The points P of \mathcal{M}^4 are called space-time points or events. Einstein's fundamental equations read as follows:¹⁰

- (i) The equation of motion for the gravitational field:

$$R_{\alpha\beta} = \kappa_G (T_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} T), \quad \alpha, \beta = 0, 1, 2, 3. \quad (0.3)$$

Here, the universal constant $\kappa_G := \frac{8\pi G}{c^4}$ depends on Newton's gravitational constant G and the velocity c of light in a vacuum.

- (ii) The equation of motion for the trajectories $x = x(\sigma)$, $\sigma_0 \leq \sigma \leq \sigma_1$, of celestial bodies (e.g., planets, the sun, stars, or galaxies) and light rays:

$$\ddot{x}^\gamma = -\dot{x}^\alpha \Gamma_{\alpha\beta}^\gamma \dot{x}^\beta, \quad \gamma = 0, 1, 2, 3. \quad (0.4)$$

This equation generalizes Newton's classical equation of motion.¹¹

Let us discuss (i) and (ii). We choose an arbitrary observer which uses the local space-time coordinates x^0, x^1, x^2, x^3 in order to describe events. The local coordinates are obtained by measurements of space positions and time. By convention, we write x instead of (x^0, x^1, x^2, x^3) . Different observers may use completely different methods for measuring space positions and time. The locality means that the real numbers x^0, x^1, x^2, x^3 do not represent coordinates for the global universe, but only for a sufficiently small spatial region and a sufficiently small time interval. The crucial change of local coordinates will be considered below. We set

- $\partial_\alpha := \frac{\partial}{\partial x^\alpha}$ (partial derivative), and
- $\dot{x}^\alpha(\sigma) := \frac{d}{d\sigma} x^\alpha(\sigma)$ (derivative with respect to the real parameter σ).

Arc length and proper time. The curve

$$C : x = x(\sigma), \quad \sigma_0 \leq \sigma \leq \sigma_1,$$

with the real parameter σ , describes a family of events, for example, the motion of a planet or the motion of a light ray. The length of the curve C is given by the integral

$$l(C) := \int_{\sigma_0}^{\sigma_1} \sqrt{g_{\alpha\beta}(x(\sigma)) \dot{x}^\alpha(\sigma) \dot{x}^\beta(\sigma)} d\sigma. \quad (0.5)$$

Here, the functions $x \mapsto g_{\alpha\beta}(x)$ are called the components of the metric tensor with respect to the local coordinates x^0, x^1, x^2, x^3 . For the motion of a planet (resp. light ray), we get $l(C) > 0$ (resp. $l(C) = 0$). The length of the curve $l(C)$ does not depend on the choice of the local coordinate system. If the trajectory $x = x(\sigma)$, $\sigma_0 \leq \sigma \leq \sigma_1$, describes the motion of a spaceship, then $l(C)/c$ is the proper time of the flight, which is measured by the crew in the spaceship during the flight.

We want to show that Einstein's equation (0.3) represents an equation for computing the components $g_{\alpha\beta}$ of the metric tensor which govern the measurement of spatial distances and proper times. We postulate that

¹⁰ We will use the Einstein summation convention, that is, we sum over equal upper and lower Greek indices from 0 to 3.

¹¹ Explicitly, this reads as $\ddot{x}^\gamma(\sigma) = -\dot{x}^\alpha(\sigma) \Gamma_{\alpha\beta}^\gamma(x(\sigma)) \dot{x}^\beta(\sigma)$.

$$g_{\alpha\beta} = g_{\beta\alpha} \quad \text{for all } \alpha, \beta = 0, 1, 2, 3.$$

In order to distinguish between the time-like coordinate x^0 and the space-like coordinates x^1, x^2, x^3 , we assume that the following definiteness conditions are always satisfied:

$$g_{00} > 0, \quad \begin{vmatrix} g_{00} & g_{01} \\ g_{10} & g_{11} \end{vmatrix} < 0, \quad \begin{vmatrix} g_{00} & g_{01} & g_{02} \\ g_{10} & g_{11} & g_{12} \\ g_{20} & g_{21} & g_{22} \end{vmatrix} > 0, \quad g := \begin{vmatrix} g_{00} & g_{01} & g_{02} & g_{03} \\ g_{10} & g_{11} & g_{12} & g_{13} \\ g_{20} & g_{21} & g_{22} & g_{23} \\ g_{30} & g_{31} & g_{32} & g_{33} \end{vmatrix} < 0. \quad (0.6)$$

In particular, these conditions are satisfied if $g_{\alpha\beta}(x) \equiv \eta_{\alpha\beta}$ for all indices $\alpha, \beta = 0, 1, 2, 3$. Here, we introduce the so-called Minkowski symbol $\eta_{\alpha\beta}$ given by

$$(\eta_{\alpha\beta}) := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (0.7)$$

In this special case, the rescaled arc length $l(C)/c$ from (0.5) is the proper time of a spaceship which moves freely, that is, the gravitational force vanishes.

The structure of the Einstein equations. Because of the symmetry properties of $g_{\alpha\beta}$, $R_{\alpha\beta}$, and $T_{\alpha\beta}$, we obtain the following:

- The Einstein equations (0.3) for the gravitational field represent a nonlinear system of 10 second-order partial differential equations for the 10 unknown functions

$$g_{00}; \quad g_{10}, g_{11}; \quad g_{20}, g_{21}, g_{22}; \quad g_{30}, g_{31}, g_{32}, g_{33}$$

which depend on the space-time variables x^0, x^1, x^2, x^3 .

- The Einstein equations (0.4) for the motion of planets and light rays represent a nonlinear system of 4 ordinary differential equations for the 4 unknown functions

$$x^\alpha = x^\alpha(\sigma), \quad \alpha = 0, 1, 2, 3,$$

which depend on the real parameter σ living in the interval $[\sigma_0, \sigma_1]$.

Changing the observer and Einstein's principle of general relativity. The following considerations are crucial for understanding the philosophy of Einstein's theory of general relativity. The Einstein equations (0.3) and (0.4) are formulated in terms of local coordinates x^0, x^1, x^2, x^3 . In terms of physics, the local coordinates describe the measurements of space positions and time positions carried out by an observer. The theory only makes sense if the following hold:

- the Einstein equations (0.3) and (0.4) are valid for all observers (i.e., for all choices of local coordinates), and
- we know the transformation laws for all the quantities under changing the local coordinates of observers:

$$\boxed{(x^0, x^1, x^2, x^3) \mapsto (x^{0'}, x^{1'}, x^{2'}, x^{3'})}. \quad (0.8)$$

To simplify notation, we briefly write $x \mapsto x'$ or $x' = x'(x)$.

(P) In terms of physics, Einstein postulated that: Physics does not depend on the choice of observers. This is Einstein's principle of general relativity.

(M) In terms of mathematics, Einstein's principle of general relativity is realized by the use of tensor calculus introduced in the second half of the 19th century.

Let us discuss this. The main points are

- the key transformation laws (0.11) and (0.12) in tensor calculus, and
- the mnemonic principle of index killing for constructing invariants.

To begin with, let us consider a typical example.

Invariance of the arc length and the proper time as a paradigm.

Naturally enough, we postulate that

The rescaled arc length (i.e., the proper time) $l(C)/c$ possesses an invariant meaning.

This means that, under a change of local coordinates (0.8), the arc length $l(C)$ remains unchanged, that is, it does not depend on the choice of the observer. Explicitly, we have

$$\begin{aligned} l(C) &= \int_{\sigma_0}^{\sigma_1} \sqrt{g_{\alpha\beta}(x(\sigma)) \dot{x}^\alpha(\sigma) \dot{x}^\beta(\sigma)} d\sigma \\ &= \int_{\sigma_0}^{\sigma_1} \sqrt{g_{\alpha'\beta'}(x'(\sigma)) \dot{x}^{\alpha'}(\sigma) \dot{x}^{\beta'}(\sigma)} d\sigma. \end{aligned} \quad (0.9)$$

Here, for the indices $\alpha = 0, 1, 2, 3$, the equation

$$x^\alpha = x^\alpha(\sigma), \quad \sigma_0 \leq \sigma \leq \sigma_1,$$

of the curve C is transformed into the equation $x^{\alpha'} = x^{\alpha'}(\sigma)$, $\sigma_0 \leq \sigma \leq \sigma_1$. Explicitly, we set $x'(\sigma) := x'(x(\sigma))$. We want to show that the transformation law

$$\boxed{g_{\alpha'\beta'}(x') = g_{\alpha\beta}(x) \cdot \frac{\partial x^\alpha(x')}{\partial x^{\alpha'}} \frac{\partial x^\beta(x')}{\partial x^{\beta'}}} \quad (0.10)$$

implies the invariance relation (0.9). Here, we sum over equal upper and lower indices from 0 to 3. In order to prove (0.9), observe that the chain rule yields

$$g_{\alpha\beta} \frac{dx^\alpha}{d\sigma} \cdot \frac{dx^\beta}{d\sigma} = g_{\alpha\beta} \frac{\partial x^\alpha}{\partial x^{\alpha'}} \frac{dx^{\alpha'}}{d\sigma} \cdot \frac{\partial x^\beta}{\partial x^{\beta'}} \frac{dx^{\beta'}}{d\sigma} = g_{\alpha'\beta'} \frac{dx^{\alpha'}}{d\sigma} \cdot \frac{dx^{\beta'}}{d\sigma}$$

along the curve C . Using the square root and integrating this over the parameter interval $[\sigma_0, \sigma_1]$, we get the claim (0.9).

Tensorial transformation laws – Ariadne's thread in tensor calculus.

The argument above is a special case of the tensor calculus which allows us to construct invariant expressions under a change of local coordinates, in a general setting. This will be thoroughly studied in Chap. 8. At this point, we would like to discuss the basic ideas. By the chain rule of calculus, we get the following two key transformation laws of tensor calculus:

- $\frac{dx^{\alpha'}}{d\sigma} = \frac{\partial x^{\alpha'}}{\partial x^\alpha} \cdot \frac{dx^\alpha}{d\sigma}$ (derivative with respect to the real parameter σ), and
- $\frac{\partial \Theta}{\partial x^{\alpha'}} = \frac{\partial \Theta}{\partial x^\alpha} \cdot \frac{\partial x^\alpha}{\partial x^{\alpha'}}$ (partial derivative)

where we sum over $\alpha = 0, 1, 2, 3$. More precisely, taking the arguments explicitly into account, this reads as

$$\frac{dx^{\alpha'}(\sigma)}{d\sigma} = \frac{\partial x^{\alpha'}(x(\sigma))}{\partial x^\alpha} \cdot \frac{dx^\alpha(\sigma)}{d\sigma} \quad (0.11)$$

and

$$\frac{\partial \Theta(x')}{\partial x^{\alpha'}} = \frac{\partial \Theta(x)}{\partial x^\alpha} \cdot \frac{\partial x^\alpha(x')}{\partial x^{\alpha'}}. \quad (0.12)$$

Here, the local coordinate x' corresponds to x , that is, $x' = x'(x)$. Moreover, we assume that the real-valued function $P \mapsto \Theta(P)$ is an invariant function on the four-dimensional space-time manifold \mathcal{M}^4 . This means that the value of Θ at the point P only depends on the event P , but not on the choice of the local coordinates which describe the event. Explicitly, $\Theta(x') = \Theta(x)$ where x' and x are related to each other by $x' = x'(x)$.¹² Let us introduce the following terminology:

- the velocity components \dot{x}^α form a *contravariant* tensorial family, and
- the partial derivatives $\partial_\alpha \Theta$ of the invariant function Θ form a *covariant* tensorial family,

In the general case, the family of functions,

$$x \mapsto T_{\beta_1 \dots \beta_n}^{\alpha_1 \dots \alpha_m}(x), \quad \alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_n = 0, 1, 2, 3,$$

is called a tensorial family of type (m, n) iff the functions are transformed like the product

$$\dot{x}^{\alpha_1} \dots \dot{x}^{\alpha_m} \partial_{\beta_1} \Theta \dots \partial_{\beta_n} \Theta$$

under a change of local coordinates. Such a tensorial family is also called m -fold contravariant and n -fold covariant. For example, by (0.10), $g_{\alpha\beta}$ transforms like the product

$$\partial_\alpha \Theta \partial_\beta \Theta.$$

Therefore, the components $g_{\alpha\beta}$ of the metric tensor form a two-fold covariant tensorial family.

The components of the Riemann curvature tensor. As in classical differential geometry, let us introduce the following Christoffel symbols:¹³

$$\Gamma_{\alpha\beta}^\gamma := \frac{1}{2}(\partial_\alpha g_{\beta\sigma} + \partial_\beta g_{\alpha\sigma} - \partial_\sigma g_{\alpha\beta})g^{\sigma\gamma}, \quad \alpha, \beta, \gamma = 0, 1, 2, 3. \quad (0.13)$$

Following Riemann, we define the components of the Riemann curvature tensor by setting¹⁴

$$R_{\alpha\beta\gamma}^\delta := \partial_\alpha \Gamma_{\beta\gamma}^\delta - \partial_\beta \Gamma_{\alpha\gamma}^\delta + \Gamma_{\alpha\mu}^\delta \Gamma_{\beta\gamma}^\mu - \Gamma_{\beta\mu}^\delta \Gamma_{\alpha\gamma}^\mu \quad (0.14)$$

where $\alpha, \beta, \kappa, \gamma, \delta = 0, 1, 2, 3$. This yields the following quantities:

- $R_{\alpha\beta\gamma\delta} := R_{\alpha\beta\gamma}^\sigma g_{\sigma\delta}$ (components of the metric Riemann curvature tensor),
- $R_{\alpha\delta} := R_{\alpha\beta\gamma\delta} g^{\beta\gamma}$ (components of the Ricci curvature tensor),
- $R := R_{\alpha\delta} g^{\alpha\delta}$ (scalar curvature – trace of the Ricci tensor),
- $G_{\alpha\beta} := R_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta}$ (components of the Einstein tensor),
- $T := T_{\alpha\beta} g^{\alpha\beta}$.

¹² For example, it turns out that the trace T of the energy-momentum tensor, and the scalar curvature R are invariant functions on the space-time manifold \mathcal{M}^4 .

¹³ By definition, the symbol $(g^{\alpha\beta})$ represents the inverse matrix to $(g_{\alpha\beta})$.

¹⁴ Mnemonically, the position of the indices κ and λ of $\Gamma_{\alpha\lambda}^\kappa$ and $R_{\alpha\beta\lambda}^\kappa$ is dictated by the symmetric formulation of the Yang matrix trick which will be discussed on page 35.

The functions $x \mapsto T_{\alpha\beta}(x)$ represent the components of the energy-momentum tensor. Furthermore, we set $T := T_{\alpha\beta}g^{\alpha\beta}$ (trace of the energy momentum tensor). In terms of physics, the energy-momentum tensor describes the distribution of mass and energy in the universe.

The gravitational force corresponds to the curvature of the four-dimensional space-time manifold. Observe the following:

- The first Einstein equation (0.3) on page 8 describes the crucial fact that the mass and energy distributions in the universe influence the curvature of the four-dimensional space-time manifold \mathcal{M}^4 . Equation (0.3) is equivalent to

$$G_{\alpha\beta} = \kappa_G T_{\alpha\beta}, \quad \alpha, \beta = 0, 1, 2, 3. \quad (0.15)$$

- As we will discuss below, the second Einstein equation (0.4) on page 8 tells us that the motion of a celestial body or a light ray corresponds to a geodesic line of the curved 4-dimensional space-time manifold \mathcal{M}^4 .

The vanishing of the gravitational force. The local vanishing of the Riemann curvature tensor corresponds to the vanishing of the gravitational force. In fact, suppose that for all indices $\alpha, \beta, \gamma, \delta = 0, 1, 2, 3$ we have

$$R_{\alpha\beta\gamma}^{\delta}(x) \equiv 0 \quad (0.16)$$

on some neighborhood of the point P_0 of the space-time manifold \mathcal{M}^4 . Using Riemann's classical argument, one can show that (0.16) implies

$$g_{\alpha\beta}(x) \equiv \eta_{\alpha\beta}, \quad \alpha, \beta = 0, 1, 2, 3 \quad (0.17)$$

on a sufficiently small neighborhood of the point P_0 . More precisely, relation (0.17) is valid after a change of local coordinates if necessary. It follows from equation (0.17) that the Christoffel symbols vanish on a sufficiently small neighborhood of P_0 . Then the Einstein equations (0.4) of motion read locally as

$$\ddot{x}^{\gamma}(\sigma) \equiv 0, \quad \gamma = 0, 1, 2, 3.$$

This equation has straight lines as solutions. For example, this corresponds to a trivial motion of space ships without any acceleration. In terms of physics, this means that the observer does not measure any gravitational force. Observe that:

If equation (0.16) is valid with respect to a fixed local coordinate system, then it is valid in every local coordinate system.

This follows from the crucial (highly nontrivial) fact that the components

$$R_{\alpha\beta\gamma}^{\delta}$$

of the Riemann curvature tensor form a tensorial family. That is, they are transformed like the product

$$\dot{x}^{\delta} \cdot \partial_{\alpha}\Theta \cdot \partial_{\beta}\Theta \cdot \partial_{\gamma}\Theta$$

under a change of local coordinates. This will be proved later on.

Einstein's local equivalence principle. Observe the following special feature. In contrast to the components $R_{\alpha\beta\gamma}^{\delta}$ of the Riemann curvature tensor, the Christoffel symbols $\Gamma_{\alpha\beta}^{\delta}$ do not form a tensorial family. For a given point P_0 of the space-time manifold \mathcal{M}^4 , it is always possible to choose a specific local coordinate system such that the Christoffel symbols vanish at the point P_0 , that is,

$$\Gamma_{\alpha\beta}^{\delta}(P_0) = 0, \quad \alpha, \beta, \delta = 0, 1, 2, 3. \quad (0.18)$$

However, as a rule, this condition is not valid in all local coordinate systems. To illustrate this by a simple example, consider an elevator which goes down with the acceleration a . If a is equal to the gravitational acceleration (i.e., $a = 9.81\text{m/s}^2$), then an observer inside the elevator does not feel anymore the gravitational field of earth. Einstein called this the local equivalence principle. This principle tells us that the gravitational force can be locally compensated by passing to an accelerated reference system. Mathematically, the local equivalence principle corresponds to (0.18). Finally, set

$$T_{\alpha\beta}^{\kappa} := \Gamma_{\alpha\beta}^{\kappa} - \Gamma_{\beta\alpha}^{\kappa}.$$

In contrast to the Christoffel symbols themselves, the so-called torsion functions $T_{\alpha\beta}^{\kappa}$ form a tensorial family. In fact, the torsion functions vanish identically, that is,

$$T_{\alpha\beta}^{\kappa} \equiv 0.$$

In other words, the Christoffel symbols are symmetric with respect to the lower indices in every local coordinate system: $\Gamma_{\alpha\beta}^{\kappa} \equiv \Gamma_{\beta\alpha}^{\kappa}$ for all indices $\alpha, \beta, \kappa = 0, 1, 2, 3$.

Dark energy. The components $T_{\alpha\beta}$ of the energy-momentum tensor allow the following decomposition:

$$T_{\alpha\beta} := T_{\alpha\beta}^{\text{class}} + T_{\alpha\beta}^{\text{CDM}} + T_{\alpha\beta}^{\text{DE}}$$

where we use the following terminology:

- $T_{\alpha\beta}^{\text{class}}$ (classical mass and energy),
- $T_{\alpha\beta}^{\text{CDM}}$ (cold dark matter),
- $T_{\alpha\beta}^{\text{DE}} = -\eta_{\text{DE}} \cdot g_{\alpha\beta}$ (dark energy),
- η_{DE} (density of dark energy),
- κ_G (universal coupling constant for gravitation),
- $\Lambda = \kappa_G \cdot \eta_{\text{DE}}$ (cosmological constant).

The quantities under consideration possess the following physical dimensions:

- $g_{\alpha\beta}$ (dimensionless),
- $R_{\alpha\beta}$ ($1/\text{length}^2$),
- $T_{\alpha\beta}$ (energy density = energy/ length^3),
- κ_G ($\text{length}/\text{energy}$),¹⁵
- Λ ($1/\text{length}^2$).

This will be studied in Volume IV.

Surprisingly enough, only 4 percent of the total mass and energy of our universe are of classical type.

Moreover, 70 percent of the total amount of energy of the universe consist of dark energy. The remaining 26 percent consist of cold dark matter.¹⁶

In Volume IV, we will use explicit solutions of the two Einstein equations (0.3) and (0.4) in order to study the following physical problems:

- the motion of the semi-axis of the planet Mercury (i.e., the slow rotation of the Perihelion of Mercury),
- the deflection of light in the gravitational field of the sun,
- the red shift in the spectrum of light caused by the gravitational field of the earth,

¹⁵ In the SI system, $\kappa_G = 2.07 \cdot 10^{-43} \text{m/J}$. This shows that the gravitational interaction is very weak compared to the scale used in daily life.

¹⁶ See G. Börner, *The Early Universe*, Springer, 2003, and S. Weinberg, *Cosmology*, Oxford University Press, 2008.

- the Big Bang, the red shift in the spectrum of distant galaxies (Hubble effect), and the accelerated expansion of the universe,
- black holes,
- the low-energy background radiation as a relict of the Big Bang.

The symmetry properties of the components of the Riemann curvature tensor. Let $\alpha, \beta, \gamma, \delta = 0, 1, 2, 3$. Then:

- $g_{\alpha\beta} = g_{\beta\alpha}$ (symmetry of the metric tensor),
- $\Gamma_{\alpha\beta}^{\delta} = \Gamma_{\beta\alpha}^{\delta}$ (symmetry of the Christoffel symbols),
- $R_{\alpha\beta} = R_{\beta\alpha}$ (symmetry of the Ricci tensor),
- $T_{\alpha\beta} = T_{\beta\alpha}$ (symmetry of the energy-momentum tensor).

Furthermore, the following hold:

- (A1) $R_{\alpha\beta\gamma}^{\delta} = -R_{\beta\alpha\gamma}^{\delta}$ (interchanging α with β).
- (A2) $R_{\alpha\beta\gamma}^{\delta} + R_{\beta\gamma\alpha}^{\delta} + R_{\gamma\alpha\beta}^{\delta} = 0$ (Ricci identity – cyclic permutation of the indices α, β, γ). This is equivalent to $R_{[\alpha\beta\gamma]}^{\delta} = 0$ (antisymmetrization with respect to α, β, γ).
- (A3) $\partial_{\mu} R_{\alpha\beta\gamma}^{\delta} + \partial_{\alpha} R_{\beta\mu\gamma}^{\delta} + \partial_{\gamma} R_{\mu\beta\gamma}^{\delta} = 0$ (Bianchi identity – cyclic permutation of the indices μ, α, β). This is equivalent to $\partial_{[\mu} R_{\alpha\beta]\gamma}^{\delta} = 0$ (antisymmetrization with respect to μ, α, β).

In order to get further symmetry properties, we consider the components of the so-called metric Riemann curvature tensor

$$R_{\alpha\beta\gamma\delta} := R_{\alpha\beta\gamma}^{\sigma} g_{\sigma\delta}$$

by lowering the upper index δ of $R_{\alpha\beta\gamma}^{\delta}$. Then:

- (B1) $R_{\alpha\beta\gamma\delta} = -R_{\beta\alpha\gamma\delta}$ (interchanging α with β),
- (B2) $R_{\alpha\beta\gamma\delta} = -R_{\alpha\beta\delta\gamma}$ (interchanging γ with δ),
- (B3) $R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta}$ (interchanging α, β with γ, δ),
- (B4) $R_{\alpha\beta\gamma\delta} + R_{\beta\gamma\alpha\delta} + R_{\gamma\alpha\beta\delta} = 0$ (metric Ricci identity – cyclic permutation of α, β, γ). This is equivalent to $R_{[\alpha\beta\gamma]\delta} = 0$ (antisymmetrization with respect to α, β, γ).
- (B5) $\partial_{\mu} R_{\alpha\beta\gamma\delta} + \partial_{\alpha} R_{\beta\mu\gamma\delta} + \partial_{\beta} R_{\mu\alpha\gamma\delta} = 0$ (metric Bianchi identity – cyclic permutation of α, β, γ). This is equivalent to $\partial_{[\mu} R_{\alpha\beta]\gamma\delta} = 0$ (antisymmetrization with respect to α, β, γ).

The relation (B4) (resp. (B5)) is obtained from (A2) (resp. (A3)) by lowering the upper index κ .

The metric Riemann curvature tensor has $4^4 = 256$ components $R_{\alpha\beta\gamma\delta}$. However, by (B1) through (B3), this large number of components is reduced to 20 independent components. Mnemonically, the symmetry properties of $R_{\alpha\beta\gamma\delta}$ motivate the definition of the components of the Ricci tensor by setting

$$R_{\alpha\delta} := R_{\alpha\beta\gamma\delta} g^{\beta\gamma}.$$

In fact, up to sign, this is the only possibility to get a nontrivial expression. In fact,

$$R_{\alpha\beta\gamma\delta} g^{\alpha\beta} = 0, \quad R_{\alpha\beta\gamma\delta} g^{\gamma\delta} = 0, \quad R_{\alpha\beta\gamma\delta} g^{\alpha\delta} = R_{\beta\alpha\delta\gamma} g^{\alpha\delta} = R_{\beta\gamma},$$

and $R_{\alpha\beta\gamma\delta} g^{\alpha\gamma} = -R_{\beta\alpha\gamma\delta} g^{\alpha\gamma} = -R_{\beta\delta}$.

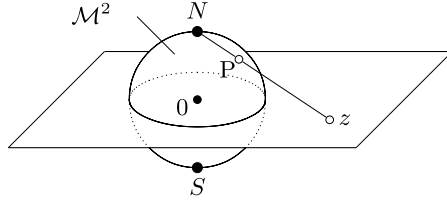


Fig. 0.1. Stereographic projection

The Riemann Curvature Tensor and the Beauty of the Gauss Theorema Egregium

For the next remarks, let us pass to the special case of a smooth 2-dimensional surface \mathcal{M}^2 which is embedded into the 3-dimensional Euclidean manifold (e.g., the surface of earth). We will use the same formulas for the components $R_{\beta\gamma\delta}^\alpha$ of the Riemann curvature tensor as introduced above, but now the indices $\alpha, \beta, \gamma, \delta$ only run from 1 to 2.

The surface theory of Gauss (1777–1855) was strongly influenced by Gauss' work as a surveyor. Under great physical pains, Gauss worked from 1821 to 1825 as a land surveyor in the kingdom of Hannover in the northern part of Germany. It almost led to his physical exhaustion. In 1822, he submitted his prize memoir “General solution of the problem of mapping parts of a given surface onto another given surface in such a way that image and pre-image become similar in their smallest parts” to the Royal Society of Sciences in Copenhagen (Denmark) for which he received the official prize. What was the importance of his work?

The mapping of surfaces onto one another, which satisfy certain given properties, is a basic problem of cartography, in particular the reproduction of parts of surfaces of the earth in plane geographic charts. Intuitively, it is impossible, for example, to map parts of the surface of the earth onto the plane and preserve the length. Therefore, one has to look for other mappings. Of great practical use are the conformal maps, that is, the angle-preserving maps. Angle preservation of geographical charts is important in navigation, that is, in determining routes of ships in charts. It turns out that conformal maps are also similar in the small. Special cases of conformal maps from the surface of the earth onto the plane are stereographic projections (see Fig. 0.1), which were already known to the Greeks, and the projection of Mercator (1512–1594) is still being used in the cartography of today. Gauss succeeded in finding a procedure to determine all conformal maps in the small for analytic surfaces.

The study of conformal maps in the large began with the Ph.D. thesis of Bernhard Riemann (1826–1866), which was written in 1851. Riemann's Ph.D. thesis contains the development of complex function theory including the famous Riemann mapping theorem. When writing his prize memoir, Gauss had apparently already worked on a more general surface theory, because he added the following Latin saying to his title page:

Ab his via sternitur ad maiora.¹⁷

The development of the general surface theory, however, was difficult, though the basic ideas were known to Gauss since 1816. On February 19, 1826, he wrote to Olbers:

¹⁷ From here the path to something more important is prepared.

I hardly know any period in my life, where I earned so little real gain for truly exhausting work, as during this winter. I found many, many beautiful things, but my work on other things has been unsuccessful for months.

Finally, on October 8, 1827, Gauss presented the general surface theory. The title of his paper was “Disquisitiones generales circa superficies curvas” (Investigations about curved surfaces). The most important result of this masterpiece in the mathematical literature is the *theorema egregium* – the egregious theorem. As a crucial quantity, Gauss introduced the *Gaussian curvature* $K(P)$ of a 2-dimensional surface \mathcal{M}^2 at the point P . For a sphere of radius r , Gauss defined

$$K(P) := \frac{1}{r^2}.$$

This tells us that the larger the radius is, the smaller is the Gaussian curvature of the sphere. By an approximation argument, Gauss generalized the curvature definition for the sphere to general surfaces \mathcal{M}^2 . In particular, the Gaussian curvature of a hyperboloid is negative (see Sect. 9.6.3). Gauss’ definition used the surrounding 3-dimensional Euclidean space. This is called an extrinsic definition. Motivated by his practical work as land surveyor, Gauss posed the following fundamental question:

Is it possible to compute the Gaussian curvature K of a 2-dimensional surface by only using measurements on the surface?

After a long fight, Gauss found that the answer is “yes”! He discovered the following sophisticated formula:

$$K(P) = \frac{R_{1221}(P)}{g(P)} \quad (0.19)$$

where $g := g_{11}g_{22} - (g_{12})^2$. This is the famous *theorema egregium*. Let us discuss this. By (0.13) and (0.14) on page 11, the following hold:

The Gaussian curvature K is an intrinsic property of the 2-dimensional surface; it depends on the components $g_{\alpha\beta}$ of the metric tensor and their first and second partial derivatives with respect to the local coordinates.

In fact, Gauss did not explicitly use the Riemann curvature tensor, but in terms of the modern terminology, his key formula can be written as (0.19).¹⁸ Concerning cartography, the *theorema egregium* tells us in rigorous terms that it is impossible to introduce geographic charts which are length preserving after rescaling. Indeed, one can show that length preserving maps preserve the components of the metric tensor. In turn, such maps preserve the Gaussian curvature. Finally, note that the Gaussian curvature of the sphere is positive, but the Gaussian curvature of the Euclidean plane vanishes.

Gauss’ theorema egregium had an enormous impact on the development of modern differential geometry and modern physics culminating in the principle “force equals curvature.” This principle is basic for both Einstein’s theory of general relativity on gravitation and the Standard Model in elementary particle physics.

¹⁸ C. Gauß, *Disquisitiones generales circa superficies curvas*, Göttinger Nachr. **6**, 99–146 (1827) (in Latin). English translation: *General Investigations of Curved Surfaces*, Raven Press, New York.

See also P. Dombrowski, 150 years after Gauss’ ‘*Disquisitiones generales circa superficies curvas*’, *Astérisque* **62** (1979).

In order to understand the intuitive meaning of both the components $R_{\alpha\beta}$ of the Ricci tensor and the scalar curvature R on the 2-dimensional surface \mathcal{M}^2 , observe that the components of the Riemann curvature tensor read as

$$R_{\alpha\beta\gamma\delta} = K(g_{\alpha\delta}g_{\beta\gamma} - g_{\alpha\gamma}g_{\beta\delta}), \quad \alpha, \beta, \gamma, \delta = 1, 2.$$

Since $g_{\alpha\beta} = g_{\beta\alpha}$, we get the following symmetry properties

$$R_{\alpha\beta\gamma\delta} = -R_{\beta\alpha\gamma\delta} = -R_{\alpha\beta\delta\gamma} = R_{\gamma\delta\alpha\beta}$$

for all indices $\alpha, \beta, \gamma, \delta = 1, 2$. Therefore, the $2^4 = 16$ components of the Riemann curvature tensor reduce to one essential component, namely,

$$R_{1221} = K(g_{11}g_{22} - g_{12}^2).$$

In fact, we have $R_{1221} = -R_{2121} = -R_{1212} = -R_{2112}$. The remaining 12 components vanish identically. For example, it follows from $R_{\alpha\beta\gamma\delta} = -R_{\beta\alpha\gamma\delta}$ that $R_{1112} = 0$. In order to simplify notation, let us introduce an orthogonal local coordinate system, that is, we have the special case where $g_{12} = g_{21} = 0$. Hence $g = g_{11}g_{22}$, and

$$g^{11} = (g_{11})^{-1}, \quad g^{22} = (g_{22})^{-1}, \quad g^{12} = g^{21} = 0.$$

This implies that:

- $R_{11} = g_{11}K$, $R_{22} = g_{22}K$, and $R_{12} = R_{21} = 0$ (Ricci tensor),
- $R = 2K$ (Ricci (or scalar) curvature).

Thus, the scalar curvature R is twice the Gaussian curvature K .

Heat conduction and the Riemann curvature tensor. Let x denote the tuple (x^1, x^2, x^3) of Cartesian coordinates. In the late 1850s, the Paris Academy posed the following problem: Find conditions such that the inhomogeneous heat conduction equation

$$\frac{\partial\Theta(x)}{\partial t} = \sum_{j,k=1}^3 g^{jk}(x) \partial_j \partial_k \Theta(x) \quad (0.20)$$

for the temperature Θ can be locally transformed into the standard heat conduction equation

$$\frac{\partial\Theta(y)}{\partial t} = \sum_{j=1}^3 \partial_j^2 \Theta(y) \quad (0.21)$$

by a change $x \mapsto y$ of local coordinates¹⁹ near the given point x_* . In terms of physics, equation (0.20) describes the heat conduction in an inhomogeneous material. If there exists such a coordinate transformation, the solution of the original complicated equation (0.20) can be reduced to the well-known solutions of the simpler equation (0.21). In 1861, Riemann solved this problem. He proved that:

The transformation is possible iff the Riemann curvature tensor R_{ijk}^l vanishes in a small neighborhood of the point x_ , that is, $R_{ijk}^l(x) \equiv 0$ for all $i, j, k, l = 1, 2, 3$.*

¹⁹ More precisely, we make the assumption that all the eigenvalues of the real symmetric (3×3) -matrix $(g^{jk}(x))$ are positive for all points $x \in \mathbb{R}^3$, and the local coordinate change $x \mapsto y$ is a local diffeomorphism on some open neighborhood of the point x_* .

In his famous 1854 lecture on the foundations of geometry, Riemann described the Riemann curvature tensor only in intuitive terms. In his 1861 paper, Riemann published the precise analytic formula of the Riemann curvature tensor for the first time.²⁰ In the textbook by M. Spivak, *A Comprehensive Introduction to Differential Geometry*, Vol. 2, Publish or Perish, Boston, one finds seven variants of the proof of Riemann's solution of the Paris Academy problem.

Riemann died in 1866 at the age of 40. His collected works fill only one volume. But his ideas, revealing deep connections between analysis, topology, and geometry, profoundly influenced the mathematics and physics of the 20th century. This is described in the beautiful book by K. Maurin, *The Riemann Legacy: Riemannian Ideas in Mathematics and Physics of the 20th Century*, Kluwer, Dordrecht.

The importance of conformal maps. Conformal mappings are essential for both classifying Riemann surfaces and proving the existence of minimal surfaces with prescribed boundary curves (the problem of Plateau (1801–1883) on soap bubbles spanned by a metallic frame).²¹

Conformal mappings play also a fundamental role in modern physics, namely, in string theory and conformal quantum field theory.

The point is that the principle of critical action in string theory is invariant under conformal mappings (which represent the gauge transformations in string theory). In 2-dimensional conformal quantum field theory, the conformal symmetry strongly restricts the structure of possible correlation functions (i.e., Green's functions). Two Riemann surfaces \mathcal{M} and \mathcal{N} are called conformally equivalent iff there exists a conformal diffeomorphism

$$\chi : \mathcal{M} \rightarrow \mathcal{N}.$$

Let $\dim_{\mathbb{R}} \mathcal{M}_g$ denote the real dimension of the space of all compact Riemann surfaces of genus g modulo conformal equivalence. By considering the description of \mathcal{M}_g by real parameters called moduli, Riemann suggested that

$$\dim_{\mathbb{R}} \mathcal{M}_g = 6g - 6 \text{ if } g = 2, 3, \dots, \quad \dim_{\mathbb{R}} \mathcal{M}_1 = \infty, \quad \dim_{\mathbb{R}} \mathcal{M}_0 = 0. \quad (0.22)$$

This was the beginning of the sophisticated theory of moduli spaces which describe the set of given geometric (or algebraic) structures up to equivalence via symmetry groups. The rigorous proof of theorem (0.22) can be given in the setting of Teichmüller spaces.²²

In what follows, we will pass back to the 4-dimensional space-time manifold \mathcal{M}^4 used in Einstein's theory of general relativity.

Two Fundamental Variational Principles

Einstein formulated the final form of the field equations (0.3) for the gravitational field in a meeting of the Prussian Academy of Sciences (Berlin) on November 25, 1915. Five days before, on November 20, 1915 in Göttingen, Hilbert lectured on an axiomatic approach based on a variational principle. A changed version of this

²⁰ B. Riemann, Mathematical remarks answering a question asked by the famous Paris Academy, pp. 391–404. In: B. Riemann, *Collected Mathematical Works*, Teubner, Leipzig, and Springer, New York, 1990.

²¹ For the solution of the Plateau problem, Ahlfors (1907–1996) was awarded the Fields medal in 1936.

²² See J. Jost, *Compact Riemann Surfaces*, Springer, 2006.

We also refer to M. Schlichenmaier, *An Introduction to Riemann Surfaces, Algebraic Curves, and Moduli Spaces*, Springer, New York, 2008.

lecture was published by Hilbert in March 1916. For example, in the special case of a universe without any matter, Hilbert's variational problem reads as

$$\int_{\mathcal{C}} R \sqrt{|g|} d^4x = \text{critical!} \quad (0.23)$$

Using the volume form $v_{\mathcal{M}^4} := \sqrt{|g|} dx^0 \wedge dx^1 \wedge dx^2 \wedge dx^3$ of the space-time manifold \mathcal{M}^4 , the variational problem (0.23) corresponds to

$$\boxed{\int_{\mathcal{C}} R \cdot v_{\mathcal{M}^4} = \text{critical!}} \quad (0.24)$$

Here, \mathcal{C} is a nonempty open subset of \mathcal{M}^4 with compact closure. The variational problem concerns all smooth metric tensors which are fixed on the boundary \mathcal{C} .

Surprisingly enough, the variational problem (0.24) is the simplest invariant variational problem related to the Riemann curvature tensor.

We will show in Volume IV that every solution of (0.24) satisfies the Euler–Lagrange equation

$$R_{\alpha\beta} = 0, \quad \alpha, \beta = 0, 1, 2, 3. \quad (0.25)$$

This coincides with the first fundamental Einstein equation

$$R_{\alpha\beta} = \kappa_G (T_{\alpha\beta} - \tfrac{1}{2} g_{\alpha\beta} T), \quad \alpha, \beta = 0, 1, 2, 3 \quad (0.26)$$

for vanishing energy-momentum tensor, $T_{\alpha\beta} \equiv 0$. The general case is obtained from (0.24) by adding the source term $\kappa_G (T_{\alpha\beta} - \tfrac{1}{2} g_{\alpha\beta} T)$.

The principle of critical arc length and geodesic lines. Consider the variational principle

$$\boxed{\int_{\sigma_0}^{\sigma_1} g_{\alpha\beta}(x(\sigma)) \dot{x}^\alpha(\sigma) \dot{x}^\beta(\sigma) d\sigma = \text{critical!}} \quad (0.27)$$

Here, we vary over all smooth curves $C : x = x(\sigma)$ with fixed initial points and fixed endpoints. The solutions of (0.27) satisfy the Euler-Lagrange equations

$$\ddot{x}^\gamma(\sigma) = -\dot{x}^\alpha(\sigma) \Gamma_{\alpha\beta}^\gamma(x(\sigma)) \dot{x}^\beta(\sigma), \quad \gamma = 0, 1, 2, 3, \quad (0.28)$$

which represent the equations of motion (0.4) for celestial bodies and light rays in general relativity. For the motion of particles which travel with a velocity smaller than that of light (e.g., the motion of planets), we can also use the variational principle of critical arc length:

$$l(C) = \int_{\sigma_0}^{\sigma_1} \sqrt{g_{\alpha\beta}(x(\sigma)) \dot{x}^\alpha(\sigma) \dot{x}^\beta(\sigma)} d\sigma = \text{critical!} \quad (0.29)$$

The solutions satisfy the equations of motion (0.28) if the parameter σ is the proper time. Einstein wrote in a letter of October 1912:

At the moment I am only concerned with the gravitational problem and I hope to overcome all the difficulties with the help of a local friend and mathematician, Marcel Grossmann (1878–1936). But it is true that, never in my life, I have worked so hard, and I am filled with a great respect for mathematics. In its subtle parts, I have regarded it, in my simplicity, as pure luxury.

The complete story of the competition between Einstein and Hilbert can be found in the newspaper article by J. Renn, Einstein, Hilbert, and the magic scrap of paper, *Frankfurter Allgemeine Zeitung*, November 20, 2005 (in German). Renn emphasizes the priority of Einstein's contributions to the creation of the theory of general relativity.

Symmetry and Klein's Invariance Principle in Geometry

Felix Klein (1849–1925) emphasized the importance of invariants in geometry.

Sophus Lie (1842–1899) discovered the importance of the linearization principle due to Newton (1643–1727) and Leibniz (1646–1716) for constructing invariants in differential geometry via Lie algebras and Lie groups.

Élie Cartan (1859–1951) combined the methods of Gauss (1777–1855) and Riemann (1826–1866) in order to describe curvature based on the ideas due to Klein and Lie.

Folklore

Klein's Erlangen program and gauge theory in physics. In the 19th century, numerous new geometries emerged in mathematics (e.g., non-Euclidean geometry and projective geometry). Missing was a general principle for classifying geometries. In 1869, the young German mathematician Felix Klein (1849–1925) and the young Norwegian mathematician Sophus Lie (1842–1899) met each other in Berlin and became close friends. Klein and Lie extensively discussed the classification problem for geometry. They agreed that symmetry groups play a distinguished role. In his 1872 Erlangen program, Felix Klein formulated the following general principle:

Geometry is the invariant theory of transformation groups.

In physics, gauge theory corresponds to a special case of this principle:

Gauge theory studies the invariants of physical fields under both space-time transformations and gauge transformations.

The main goal of gauge theory is the formulation of

- variational principles (principle of critical action) and
- partial differential equations (Euler–Lagrange equations)

which are invariant under both space-time transformations and gauge transformations. Such invariant variational principles and differential equations appear in:

- (a) electrodynamics (the Maxwell equations),
- (b) the Standard Model in elementary particle physics,
- (c) the theory of general relativity (e.g., the Standard Model in cosmology).

In this connection, our main goal is

to create a differential calculus which respects both space-time transformations and gauge transformations.

It was the beautiful idea of Élie Cartan to combine curvature in differential geometry with local symmetry. Nowadays we know that precisely this idea is basic for modern physics, too.

A glance at the history of invariant theory. Invariant theory was created in the 19th century by George Boole (1815–1864), James Sylvester (1814–1897), and Arthur Cayley (1821–1895). Hermann Weyl wrote:²³

²³ H. Weyl, *Invariants*, *Duke Math. J.* **5** (1939), 489–502.

The theory of invariants came into existence about the middle of the nineteenth century somewhat like Minerva:²⁴ a grown-up virgin, mailed in the shining armor of algebra, she sprang forth from Cayley's Jovian head. Her Athens over which she ruled and which she served as a tutelary and beneficent goddess was projective geometry.

Cayley was a master in doing long computations and in inventing algorithms. A brief history of invariant theory can be found in the introduction of Peter Olver's book: *Classical Invariant Theory*, Cambridge University Press, 1999. We also refer to Felix Klein's famous book: *Development of Mathematics in the 19th Century*, Math. Sci. Press, New York, 1979.

The goal of invariant theory. We are given a mathematical object \mathcal{O} and a symmetry group \mathcal{G} which transforms the object \mathcal{O} . The final goal is to construct \mathcal{G} -invariants of \mathcal{O} . That is, we are looking for quantities which are assigned to \mathcal{O} and which are invariant under the action of the symmetry group \mathcal{G} . Moreover, we are interested in determining a complete system of invariants. By definition, a system of \mathcal{G} invariants of \mathcal{O} is called complete iff it uniquely determines the object \mathcal{O} up to symmetry operations contained in the group \mathcal{G} .

A typical example. Consider the quadratic equation

$$ax^2 + 2bxy + dy^2 = 1, \quad (x, y) \in \mathbb{R}^2 \quad (0.30)$$

with the real coefficients a, b, d . The theorem of principal axes tells us that there exists a rotation $(x, y) \mapsto (\xi, \eta)$ such that equation (0.30) is transformed into

$$\alpha\xi^2 + \beta\eta^2 = 1, \quad (\xi, \eta) \in \mathbb{R}^2 \quad (0.31)$$

with the real coefficients α and β . Moreover, we have the invariants

$$ad - b^2 = \alpha\beta \quad \text{and} \quad a + d = \alpha + \beta.$$

This immediately implies the following two statements:

- (i) Ellipse: If $ad - b^2 > 0$ and $a + d > 0$, then $\alpha > 0$ and $\beta > 0$. Thus, equation (0.30) represents an ellipse.
- (ii) Hyperbola: If $ad - b^2 < 0$, then α and β have different signs. For example, $\alpha > 0$ and $\beta < 0$. Thus, equation (0.30) represents a hyperbola.

The point is that the matrix

$$A := \begin{pmatrix} a & b \\ b & d \end{pmatrix}$$

has the eigenvalues α, β . This means that the equation

$$\det(A - \lambda I) = \lambda - \operatorname{tr}(A)\lambda + \det(A) = 0 \quad (0.32)$$

has the zeros α and β . This will be studied in Sect. 3.8.1 on page 200 (theorem of principal axes). Equation (0.32) was used by Lagrange (1736–1813) in order to compute the long-time (secular) perturbations of the orbit of a planet under the influence of the other planets. Therefore, this equation is called the secular equation. Gauss used the two invariants $\det(A)$ and $\operatorname{tr}(A)$ in order to define the Gaussian curvature and the mean curvature of a 2-dimensional surface, respectively (see Sect. 9.6.3 on page 628). James Sylvester (1814–1897) said in 1864:²⁵

²⁴ Minerva was the ancient Roman goddess of wisdom and the art, identified with the Greek goddess Athena.

²⁵ J. Sylvester, *Collected Mathematical Papers*, Vol. 2, p. 380, Cambridge University Press, 1864.

As all roads lead to Rome so I find in my own case at least that all algebraic inquiries, sooner or later, end at the Capitol of modern algebra over whose shining portal is inscribed the Theory of Invariants.

Invariant theory is essential for modern physics. In the present volume we will encounter invariant theory again and again.

Einstein's Principle of General Relativity and Invariants – the Geometrization of Physics

Einstein emphasized the importance of invariants in physics.
Folklore

The components $R^\alpha_{\beta\gamma\delta}$ of the Riemann curvature tensor depend on the choice of local space-time coordinates x^0, x^1, x^2, x^3 , that is, they depend on the choice of the observer. Recall that Einstein's principle of general relativity tells us that:

Physics is independent of the choice of the observer.

This means that proper physical quantities have to be independent of the choice of local coordinates. As an example, choose two events P_0 and P_1 (e.g., the depart P_0 and the return P_1 of a space ship to earth). Consider the difference

$$\Delta t := \frac{x^0(P_1)}{c} - \frac{x^0(P_0)}{c}$$

where $x^\alpha(P)$, $\alpha = 0, 1, 2, 3$, denotes the local coordinates of the event P . The quantity Δt has the physical dimension of time. But, as a rule, Δt has not an immediate physical meaning because it depends on the choice of the local coordinates for describing the measurements. In contrast to this, the proper time $l(C)/c$ (i.e., the rescaled arc length) introduced by (0.5) on page 8 does not depend on the choice of local coordinates, and hence it possesses an invariant meaning called the proper time interval which can be measured by physical experiments. We refer to the twin paradox considered in Sect. 18.4.3 on page 926.

In the theory of general relativity, transformations of local space-time coordinates are called gauge transformations. Using this term, one can say that

Einstein wanted to construct his theory of general relativity in such a way that it is gauge invariant.

In other words, starting with his philosophical principle of general relativity, Einstein was looking for a mathematical approach which describes invariants in terms of local coordinates. The prototype of such an approach is given by cartography.

Cartography as a paradigm. In cartography, parts of the surface of earth are described by local geographic charts collected in a geographic atlas. The Euclidean coordinates of each chart are called local coordinates of earth. Obviously, geometric properties of the surface of earth do not depend on the choice of the geographic charts, for example, the distance of two points on the surface of earth does not depend on the choice of local coordinates. Geometric properties are invariants with respect to the possible choices of local coordinates.

Intuitively spoken, Einstein looked for higher-dimensional cartography.

His friend – the mathematician Marcel Grossmann (1878–1936) – told him that Riemann generalized Gauss' theory of cartography to higher dimensions and that there exists a well-developed calculus for higher-dimensional manifolds, namely, the Ricci calculus due to Gregorio Ricci-Curbastro (1853–1925). By the help of

Grossmann, Einstein studied the Ricci calculus and he applied it to his theory of gravitation.

The geometrization of physics. Geometry is a mathematical model for describing both invariant geometric properties and their representation by local coordinates. In ancient times, one only considered invariant geometric properties. The description of geometric properties by coordinates dates back to René Descartes (1596–1650). In 1667 Descartes published his “Discours de la méthode” which contains, among a detailed philosophical investigation and its application to the sciences, the foundation of analytic geometry (e.g., the use of Cartesian coordinates).²⁶

Einstein geometrized gravitation in his 1915 theory of general relativity. Quantum mechanics was geometrized by Dirac, as a unitary geometry of Hilbert spaces. In the introduction to his book “The Principles of Quantum Mechanics,” Clarendon Press, Oxford, 1930, the young Dirac (1902–1984) wrote:

The important things in the world appear as invariants . . . The things we are immediately aware of are the relations of these invariants to a certain frame of reference . . . The growth of the use of transformation theory, as applied first to relativity and later to the quantum theory, is the essence of the new method in theoretical physics.

Finally, note that the Standard Model in particle physics starts from a classical field theory which is closely related to the geometry of specific fiber bundles.

Invariant formulation of the fundamental Einstein equations. To begin with, let us introduce the following notation:

- $\mathbf{v} = v^\alpha \partial_\alpha$ (intrinsic tangent vector),
- $\mathbf{g} := g_{\alpha\beta} dx^\alpha \otimes dx^\beta$ (metric tensor field),
- $\langle \mathbf{u} | \mathbf{v} \rangle := \mathbf{g}(\mathbf{u}, \mathbf{v})$ (indefinite Hilbert inner product), $\langle \mathbf{u} | \mathbf{v} \rangle = g_{\alpha\beta} u^\alpha v^\beta$,
- $\text{Ric}(\mathbf{g}) = R_{\alpha\beta} dx^\alpha \otimes dx^\beta$ (Ricci tensor field), $R_{\alpha\beta} = R_{\alpha\kappa\lambda\beta} g^{\kappa\lambda}$,
- $R = R_{\alpha\beta} g^{\alpha\beta}$ (scalar curvature or Ricci curvature – trace of the Ricci tensor),
- $\mathbf{T} = T_{\alpha\beta} dx^\alpha \otimes dx^\beta$ (energy-momentum tensor field),
- $\text{tr}(\mathbf{T}) = T_{\alpha\beta} g^{\alpha\beta}$ (trace of the energy-momentum tensor field), $\text{tr}(\mathbf{T}) \equiv T^\alpha_\alpha$,
- $\mathbf{G} = \text{Ric}(\mathbf{g}) - \frac{1}{2} R \mathbf{g}$ (the Einstein tensor field).

Then, in general relativity, the two fundamental Einstein equations (0.3) and (0.4) on page 8 read as follows in an invariant way:

- (i) The equation of motion for the gravitational field:

$$\boxed{\text{Ric}(\mathbf{g}) = \kappa_G (\mathbf{T} - \frac{1}{2} \text{tr}(\mathbf{T}) \cdot \mathbf{g})}. \quad (0.33)$$

This is equivalent to: $\mathbf{G} = \kappa_G \mathbf{T}$.

- (ii) The equation of motion for the trajectories $x = x(\sigma)$, $\sigma_0 \leq \sigma \leq \sigma_1$, of celestial bodies (e.g., planets, the sun, stars, or galaxies) and light rays:

$$\boxed{\frac{D\dot{x}(\sigma)}{d\sigma} = 0, \quad \sigma_0 \leq \sigma \leq \sigma_1}. \quad (0.34)$$

Let us discuss (i).²⁷ To begin with, fix the event P . Choose the local coordinates $x = (x^0, x^1, x^2, x^3)$. Assume that the event P has the local coordinate x_P . Let us

²⁶ Descartes’ Latin name was Cartesius.

²⁷ The covariant derivative $\frac{D\dot{x}(\sigma)}{d\sigma}$ will be discussed on page 43. The invariance of the two fundamental equations (0.33) and (0.34) follows from the principle of killing indices to be discussed on page 29.

start with the trajectory $x = x(\sigma)$ which passes through the point P , that is, $x(0) = x_P$. Set $v^\alpha := \dot{x}^\alpha(0)$ if $\alpha = 0, 1, 2, 3$. Moreover, we define the differential operator

$$\mathbf{v} := v^\alpha \partial_\alpha.$$

We want to show that this is an invariant notion, that is,

$$\boxed{\mathbf{v} = v^\alpha \partial_\alpha = v^{\alpha'} \partial_{\alpha'}} \quad (0.35)$$

with the tensorial transformation laws $v^{\alpha'} = \frac{\partial x^{\alpha'}(x)}{\partial x^\alpha} v^\alpha$ and $\partial_{\alpha'} = \frac{\partial x^\alpha(x')}{\partial x^{\alpha'}} \partial_\alpha$.

To this end, let us consider different local coordinates $x^{0'}, x^{1'}, x^{2'}, x^{3'}$ given by the transformation $x' = x'(x)$ together with the inverse transformation $x = x(x')$. Then the curve $x = x(\sigma)$ corresponds to $x' = x'(x(\sigma))$. Let $\Theta = \Theta(P)$ be a given invariant real-valued function, that is,

$$\Theta(x(\sigma)) = \Theta(x'(x(\sigma))).$$

Differentiating this with respect to the real parameter σ at the value $\sigma = 0$, we get

$$\frac{\partial \Theta}{\partial x^\alpha} \frac{dx^\alpha(0)}{d\sigma} = \frac{\partial \Theta}{\partial x^\beta} \frac{\partial x^\beta}{\partial x^{\alpha'}} \frac{dx^{\alpha'}(0)}{d\sigma} = \frac{\partial \Theta}{\partial x^{\alpha'}} \frac{dx^{\alpha'}(0)}{d\sigma},$$

by the chain rule. Hence

$$\partial_\alpha \Theta \cdot v^\alpha = \partial_\alpha \Theta \cdot \dot{x}^\alpha(0) = \partial_{\alpha'} \Theta \cdot \dot{x}^{\alpha'}(0) = \partial_{\alpha'} \Theta \cdot v^{\alpha'}.$$

This yields $(v^\alpha \partial_\alpha) \Theta = (v^{\alpha'} \partial_{\alpha'}) \Theta$ for all smooth functions Θ . This is the claim (0.35).

Lie's intrinsic tangent vectors on the space-time manifold \mathcal{M}^4 . Note the following:

By definition, tangent vectors of the 4-dimensional space-time manifold \mathcal{M}^4 are linear differential operators of first order with constant coefficients.

Moreover, smooth tangent vector fields on \mathcal{M} are linear differential operators of first order with smooth coefficient functions. This definition dates back to the work of Lie in the second half of the 19th century. For a moment, the definition sounds strange. Let us discuss this. Following Gauss, we have to distinguish between the extrinsic and the intrinsic approach to differential geometry. To illustrate this, consider a 2-dimensional sphere \mathcal{M}^2 embedded in the 3-dimensional Euclidean manifold (e.g., the surface of earth).

- **Extrinsic tangent vectors:** Intuitively, the tangent plane $T_P \mathcal{M}^2$ of the sphere \mathcal{M}^2 at the point P is a 2-dimensional plane in the 3-dimensional Euclidean manifold. This plane is orthogonal to the normal vector of the sphere \mathcal{M}^2 at the point P . In this setting, the definition of $T_P \mathcal{M}^2$ is based on the surrounding Euclidean manifold \mathbb{E}^3 . Such a definition is called an extrinsic one. Extrinsic tangent vectors of the sphere are precisely the position vectors of the Euclidean manifold \mathbb{E}^3 at the point P which are orthogonal to the external normal unit vector of the sphere at the point P .
- **Intrinsic tangent vectors:** We will show on page 529 that there exists a natural linear isomorphism

$$T_P \mathcal{M}^2 \simeq D_P \mathcal{M}^2 \quad (0.36)$$

between the linear space $T_P\mathcal{M}^2$ of extrinsic tangent vectors and the linear space $D_P\mathcal{M}^2$ of linear differential operators (of first order with constant coefficients) on the sphere \mathcal{M}^2 at the point P . Intuitively, these differential operators act on temperature fields $\Theta : \mathcal{M}^2 \rightarrow \mathbb{R}$ on the sphere. The notion of a linear differential operator (of first order) on the sphere does not use the surrounding Euclidean manifold. Therefore, motivated by (0.36), linear differential operators contained in $D_P\mathcal{M}^2$ are called intrinsic tangent vectors.

In the case of a sphere in the 3-dimensional Euclidean manifold, the close relation between extrinsic tangent vectors, the directional derivative, intrinsic tangent vectors, and derivations is discussed in Sect. 8.15 on page 529. Motivated by velocity vectors of fluids on earth (e.g., rivers and oceans), tangent vectors are also called velocity vectors.

In the case of the 4-dimensional space-time manifold \mathcal{M}^4 , we do not want to use any surrounding space. Therefore, we intrinsically describe tangent vectors by linear differential operators. The symbol $T_P\mathcal{M}^4$ denotes the set of all tangent vectors of the space-time manifold \mathcal{M}^4 at the point P . This is called the tangent space of \mathcal{M}^4 at the point P . Naturally enough, suppose that $\Theta : \mathcal{M}^4 \rightarrow \mathbb{R}$ is a real-valued function on the space-time manifold \mathcal{M}^4 . Then we define

$$\mathbf{v}_P(\Theta) := v^\alpha(x) \partial_\alpha \Theta(x)$$

where x denotes the local coordinate of P . This definition does not depend on the choice of local coordinates. The linear differential operator

$$\mathbf{v}_P := v^\alpha(x) \partial_\alpha$$

is called a tangent vector of the space-time manifold \mathcal{M}^4 at the point P .

Élie Cartan's algebraization of calculus and infinitesimals. In a heuristic manner, Newton (1643–1727) and Leibniz (1646–1716) used “infinitesimally small” quantities possessing the typical properties $dx > 0$ and “ $dx^2 = 0$.” In fact, there are *no* real numbers which possess such strange properties.²⁸ Observe that:

In modern differential geometry, differentials like dx^α are well-defined mathematical objects, namely, linear functionals on the tangent space.

That is, $dx^\alpha \in T_P^*(\mathcal{M}^4)$. Let us discuss this. If $\mathbf{v} = v^\alpha \partial_\alpha$ is an element of the tangent space $T_P\mathcal{M}^4$, then we define²⁹

$$dx^\alpha(\mathbf{v}) := v^\alpha.$$

Thus, the linear functional dx^α assigns to the (abstract) tangent vector \mathbf{v} the real value v^α measured by physical experiment. The linear functionals

$$\omega : T_P\mathcal{M}^4 \rightarrow \mathbb{R}$$

are precisely given by the linear combinations $\omega = a_\alpha dx^\alpha$ where the symbols a_0, a_1, a_2, a_3 are fixed, but otherwise arbitrary real numbers. These linear functionals form the dual space to the tangent space $T_P\mathcal{M}^4$ which is called the cotangent space $T_P^*\mathcal{M}^4$ of the space-time manifold \mathcal{M}^4 at the point P . Tensor products will

²⁸ In non-standard analysis, one rigorously introduces infinitesimally small numbers ι which are contained in a field extension $^*\mathbb{R}$ of the classical field \mathbb{R} of real numbers, and which have the property that $0 < \iota < \varepsilon$ for all positive real numbers ε . In addition, $\iota^2 > 0$ (see Sect. 4.6 of Vol. II).

²⁹ Note that dx^α , \mathbf{v} , and v^α depend on the choice of the point P .

be thoroughly studied in Sect. 2.1.2. For example, the symbol $dx^\alpha \otimes dx^\beta$ denotes a bilinear map from $T_P \mathcal{M}^4 \times T_P \mathcal{M}^4$ to \mathbb{R} given by

$$(dx^\alpha \otimes dx^\beta)(\mathbf{u}, \mathbf{v}) := dx^\alpha(\mathbf{u}) \cdot dx^\beta(\mathbf{v}) = u^\alpha v^\beta$$

for all tangent vectors $\mathbf{u}, \mathbf{v} \in T_P(\mathcal{M}^4)$. Élie Cartan introduced the antisymmetric wedge product of differentials by setting

$$dx^\alpha \wedge dx^\beta := dx^\alpha \otimes dx^\beta - dx^\beta \otimes dx^\alpha.$$

Hence

$$(dx^\alpha \wedge dx^\beta)(\mathbf{u}, \mathbf{v}) = dx^\alpha(\mathbf{u}) dx^\beta(\mathbf{v}) - dx^\alpha(\mathbf{v}) dx^\beta(\mathbf{u}) = u^\alpha v^\beta - v^\alpha u^\beta.$$

In particular, we obtain

$$dx^\alpha \wedge dx^\alpha = 0, \quad \alpha = 0, 1, 2, 3, \quad (0.37)$$

which replaces the heuristic relation “ $(dx^\alpha)^2 = 0$ ” used by Newton and Leibniz.

Gauge Transformations

Some important gauge groups. For the convenience of the reader, let us start with summarizing some matrix notation which will be used again and again in this volume. Let $N = 1, 2, \dots$

- The group $GL(N, \mathbb{C})$ (resp. $GL(N, \mathbb{R})$) consists of all complex (resp. real) invertible $(N \times N)$ -matrices. Furthermore, we have the following chain of subgroups: $SU(N) \subseteq U(N) \subseteq GL(N, \mathbb{C})$.
- $G \in U(N)$ iff $G \in GL(N, \mathbb{C})$ and $G^{-1} = G^\dagger$.
- $G \in SU(N)$ iff $G \in U(N)$ and $\det G = 1$.

In particular, the group $U(1)$ consists of all complex numbers z with $|z| = 1$.

Einstein's gauge transformations in the theory of general relativity. Suppose that the event P corresponds to the local coordinates $x = (x^0, x^1, x^2, x^3)$ and $x' = (x^{0'}, x^{1'}, x^{2'}, x^{3'})$. As a preparation, let us introduce the following matrices:

$$v = \begin{pmatrix} v^0 \\ v^1 \\ v^2 \\ v^3 \end{pmatrix}, \quad v' = \begin{pmatrix} v^{0'} \\ v^{1'} \\ v^{2'} \\ v^{3'} \end{pmatrix}, \quad \partial\Theta = \begin{pmatrix} \partial_0\Theta \\ \partial_1\Theta \\ \partial_2\Theta \\ \partial_3\Theta \end{pmatrix}, \quad \partial'\Theta = \begin{pmatrix} \partial_{0'}\Theta \\ \partial_{1'}\Theta \\ \partial_{2'}\Theta \\ \partial_{3'}\Theta \end{pmatrix}, \quad (0.38)$$

$$G(P) = \begin{pmatrix} G_0^{0'} & G_1^{0'} & G_2^{0'} & G_3^{0'} \\ G_0^{1'} & G_1^{1'} & G_2^{1'} & G_3^{1'} \\ G_0^{2'} & G_1^{2'} & G_2^{2'} & G_3^{2'} \\ G_0^{3'} & G_1^{3'} & G_2^{3'} & G_3^{3'} \end{pmatrix}, \quad G(P)^{-1} = \begin{pmatrix} G_{0'}^0 & G_{1'}^0 & G_{2'}^0 & G_{3'}^0 \\ G_{0'}^1 & G_{1'}^1 & G_{2'}^1 & G_{3'}^1 \\ G_{0'}^2 & G_{1'}^2 & G_{2'}^2 & G_{3'}^2 \\ G_{0'}^3 & G_{1'}^3 & G_{2'}^3 & G_{3'}^3 \end{pmatrix}. \quad (0.39)$$

Here, we set $G_\alpha^{\alpha'} := \frac{\partial x^{\alpha'}(x)}{\partial x^\alpha}$ and $G_{\alpha'}^\alpha := \frac{\partial x^\alpha(x')}{\partial x^{\alpha'}}$. By the chain rule,

$$G_{\alpha}^{\alpha'} G_{\beta'}^{\alpha} = \frac{\partial x^{\alpha'}}{\partial x^{\alpha}} \frac{\partial x^{\alpha}}{\partial x^{\beta'}} = \frac{\partial x^{\alpha'}}{\partial x^{\beta'}} = \delta_{\beta'}^{\alpha'}.$$

Therefore, the matrix $(G_{\alpha}^{\alpha'})$ is the inverse matrix to $G(P) := (G_{\alpha}^{\alpha'})$. By (0.35), we get $v^{\alpha'} = G_{\alpha}^{\alpha'} v^{\alpha}$. Hence

$$\boxed{v' = G(P)v.} \quad (0.40)$$

This is called the gauge transformation of the components of the tangent vector $\mathbf{v} = v^{\alpha} \partial_{\alpha} = v^{\alpha'} \partial_{\alpha'}$. The matrix $G(P)$ is contained in the Lie group $GL(4, \mathbb{R})$. The transformation formula $\partial_{\alpha'} \Theta = \frac{\partial x^{\alpha}}{\partial x^{\alpha'}} \partial_{\alpha} \Theta = G_{\alpha'}^{\alpha} \partial_{\alpha} \Theta$ reads as

$$\boxed{(\partial' \Theta)^d = (\partial \Theta)^d G(P)^{-1}.} \quad (0.41)$$

For example, this implies the invariance relation (0.35). In fact, we get

$$(\partial' \Theta)^d v' = (\partial \Theta)^d G(P)^{-1} G(P) v = (\partial \Theta)^d v = \mathbf{v}(\Theta).$$

Einstein's gauge transformations in the special theory of relativity. In the theory of special relativity, gauge transformations correspond to a change of inertial systems. Then the matrix $G(P)$ from (0.40) represents a Lorentz transformation (see Chap. 18).

Dirac's gauge transformations in quantum mechanics. We want to discuss Dirac's quotation mentioned on page 23. To this end, let X be a complex n -dimensional Hilbert space. The unit vectors ψ of X are called physical states. Choose a complete orthonormal system e_1, \dots, e_n of X . Then the Fourier expansion reads as

$$\psi = \sum_{j=1}^n \langle e_j | \psi \rangle e_j.$$

In this setting, we have to distinguish between

- the (invariant) physical state ψ , and
- the local coordinates $\langle e_j | \psi \rangle, j = 1, \dots, n$, of ψ (also called the Feynman probability amplitudes of ψ).

In terms of physics, the choice of the orthonormal basis e_1, \dots, e_n corresponds to a measurement device. If the given quantum particle is in the state ψ , then the real number

$$|\langle \psi | e_j \rangle|^2$$

is the probability for measuring the particle in the state e_j of the measurement device.

In Dirac's setting of quantum mechanics, a gauge transformation corresponds to a change of the measurement device.

That is, in terms of mathematics, we pass from the orthonormal basis e_1, \dots, e_n to the orthonormal basis e_1', \dots, e_n' . This basis change can be described by a unitary transformation³⁰

$$G: X \rightarrow X$$

³⁰ Recall that the linear operator $G: X \rightarrow X$ is called unitary iff it is bijective and it preserves inner products. The group of all unitary transformations of the Hilbert space X is denoted by $U(X)$ (unitary group of X).

defined by $Ge_j := e_j$ for $j = 1, \dots, n$. In this sense, quantum mechanics corresponds to the unitary geometry of the Hilbert space X . By definition, this is the theory of invariants under the unitary transformation group $U(X)$. For example, the transition amplitude $\langle \psi | \varphi \rangle$ is a unitary invariant. We also define

$$dx^j(\psi) := \langle e_j | \psi \rangle, \quad j = 1, \dots, n.$$

Here, the linear functional $dx^j : X \rightarrow \mathbb{C}$ assigns to the physical state ψ the local coordinate $\langle e_j | \psi \rangle$ (probability amplitude) which depends on the choice of the measurement device.

This argument can be immediately generalized to complex infinite-dimensional separable Hilbert spaces. Such spaces possess a countable orthonormal basis e_1, e_2, \dots (also called complete orthonormal system).

In the theory of special relativity, gauge transformations correspond to a change of inertial systems, which corresponds to a change of a pseudo-orthonormal basis of the Minkowski space which is an indefinite Hilbert space (see Chap. 18).

Yang's gauge transformations via local phase factors. We are given the function $\psi : \mathbb{R}^4 \rightarrow \mathbb{C}$. Consider the transformation formula

$$\psi_+(x) = G(x)\psi(x), \quad x \in \mathbb{R}^4 \quad (0.42)$$

where $G(x)$ is an element of the Lie group $U(1)$. Explicitly,

$$G(x) = e^{i\chi(x)}, \quad x \in \mathbb{R}^4.$$

Here, the so-called phase $\chi(x)$ is a real number which depends on the choice of the space-time point x in \mathbb{R}^4 . Therefore, $G(x)$ is called a local phase factor according to Yang. In terms of physics, the function ψ is the wave function of an electron, and the map $\psi(x) \mapsto \psi_+(x)$ given by (0.42) is called a gauge transformation. In the Standard Model in particle physics, this situation is generalized in the following way:

- The function $\psi : \mathbb{R}^4 \rightarrow \mathbb{C}^N$ describes the basic particles (i.e., quarks and leptons),
- and the local phase factor $G(x)$ is an element of the Lie group $GL(N, \mathbb{C})$.

More precisely, we have $G(x) \in \mathcal{G}$ for all $x \in \mathbb{R}^4$. Here, \mathcal{G} is a subgroup of $GL(N, \mathbb{C})$ which is isomorphic to the Lie group $U(1) \times SU(2) \times SU(3)$. The latter group is called the gauge group of the Standard Model in particle physics. Summarizing, we will encounter the following groups as gauge groups:

- $SO(1, 3)$ – Lorentz group (Einstein's theory of special relativity),
- $U(1)$ (Maxwell's theory of electromagnetism),
- $SU(2)$ (the original Yang–Mills theory for the local isospin phase factor),
- $U(1) \times SU(2)$ (electroweak interaction),
- $SU(3)$ (strong interaction – quantum chromodynamics),
- $U(1) \times SU(2) \times SU(3)$ (Standard Model in particle physics – combining electroweak interaction with strong interaction).

Observe that the gauge groups $U(1)$, $SU(2)$, and $SU(3)$, as well as their direct product $U(1) \times SU(2) \times SU(3)$ are compact Lie groups, whereas the gauge group $SO(1, 3)$ is not compact, but only locally compact. The representations of compact groups are much easier to handle than the representations of non-compact, locally compact Lie groups.

Furthermore, note the following. Let $\text{Diff}(\mathcal{M}^4)$ denote the group of all diffeomorphisms $\chi : \mathcal{M}^4 \rightarrow \mathcal{M}^4$ of the space-time manifold \mathcal{M}^4 onto itself.

- According to Einstein's principle of general relativity, physical quantities have to be invariant under the group $\text{Diff}(\mathcal{M}^4)$. In contrast to the finite-dimensional Lie groups $U(1), SU(2), SU(3), SO(1, 3)$, the group $\text{Diff}(\mathcal{M}^4)$ is an infinite-dimensional generalized Lie group.
- String theory is based on conformal symmetry (like the theory of minimal surfaces, the theory of Riemann surfaces, and the conformal quantum field theory). On an infinitesimal level, this is described by an infinite-dimensional Lie algebra called the Virasoro algebra.

Therefore, the theory of finite-dimensional and infinite-dimensional groups (resp. Lie algebras) and their invariants play a fundamental role in modern physics. Important contributions to this topic were made by Élie Cartan, Weyl (compact Lie groups), Wigner, Bargmann, Gelfand, and Harish-Chandra (noncompact groups), and Victor Kac (infinite-dimensional Lie algebras).

Construction of Invariants by the Principle of Killing Indices

Mnemonically, the principle of killing indices works on its own.
Folklore

Fix $n, m = 1, 2, \dots$. In what follows we will sum over equal upper and lower Greek indices from 0 to 3. Let

$$T_{\beta_1 \dots \beta_n}^{\alpha_1 \dots \alpha_m}, \quad S_{\alpha_1 \dots \alpha_m}^{\beta_1 \dots \beta_n},$$

and $U_{\delta_1 \dots \delta_s; \beta_1 \dots \beta_n}^{\gamma_1 \dots \gamma_r; \alpha_1 \dots \alpha_m}$ be tensorial families on the 4-dimensional space-time manifold \mathcal{M}^4 . Set

$$V_{\delta_1 \dots \delta_s}^{\gamma_1 \dots \gamma_r} := U_{\delta_1 \dots \delta_s; \beta_1 \dots \beta_n}^{\gamma_1 \dots \gamma_r; \alpha_1 \dots \alpha_m} S_{\alpha_1 \dots \alpha_m}^{\beta_1 \dots \beta_n}.$$

Then, we have the following three very useful principles for constructing invariants:

- (K1) $T_{\beta_1 \dots \beta_n}^{\alpha_1 \dots \alpha_m} S_{\alpha_1 \dots \alpha_m}^{\beta_1 \dots \beta_n}$ is an invariant function on \mathcal{M}^4 .
- (K2) $V_{\delta_1 \dots \delta_s}^{\gamma_1 \dots \gamma_r}$ is a tensorial family (r -fold contravariant and s -fold covariant).
- (K3) $T_{\alpha_1 \dots \alpha_n}^{\beta_1 \dots \beta_m} dx^{\alpha_1} \otimes \dots \otimes dx^{\alpha_n} \otimes \partial_{\beta_1} \otimes \dots \otimes \partial_{\beta_m}$ is an invariant mathematical object denoted by \mathbf{T} .

Let us explain the meaning of the mathematical object \mathbf{T} .

- (i) Consider first the special case where

$$\mathbf{T} := T_{\gamma}^{\alpha\beta} dx^{\gamma} \otimes \partial_{\alpha} \otimes \partial_{\beta}.$$

Fixing the point P of the manifold \mathcal{M}^4 , we get

$$\mathbf{T}_P := T_{\gamma}^{\alpha\beta}(P) dx^{\gamma} \otimes \partial_{\alpha} \otimes \partial_{\beta}.$$

As usual, the tensor product $\partial_{\alpha} \otimes \partial_{\beta}$ of two differential operators acts on smooth functions $(x, y) \mapsto f(x, y)$ by setting

$$(\partial_{\alpha} \otimes \partial_{\beta})f(x, y) = \frac{\partial}{\partial x^{\alpha}} \frac{\partial}{\partial y^{\beta}} f(x, y) = \frac{\partial^2 f(x, y)}{\partial x^{\alpha} \partial y^{\beta}}.$$

Here, $x = (x^0, x^1, x^2, x^3)$ and $y = (y^0, y^1, y^2, y^3)$. Now choose a tangent vector $\mathbf{u} := u^{\sigma}(P)\partial_{\sigma}$ of \mathcal{M}^4 at the point P . Then

$$\mathbf{T}_P(\mathbf{u}) = T_{\gamma}^{\alpha\beta}(P) dx^{\gamma}(\mathbf{u}) \otimes \partial_{\alpha} \otimes \partial_{\beta} = T_{\gamma}^{\alpha\beta}(P) u^{\gamma}(P) \partial_{\alpha} \otimes \partial_{\beta}.$$

Finally, we obtain

$$\mathbf{T}_P(\mathbf{u}) = T_\gamma^{\alpha\beta}(P) u^\gamma(P) \frac{\partial^2}{\partial x^\alpha \partial y^\beta}.$$

This is a linear differential operator of second order with coefficients which depend on the point P .

- (ii) Consider now the general case of (K3). For all tangent vectors $\mathbf{u}_1, \dots, \mathbf{u}_n$ of the manifold \mathcal{M}^4 at the fixed point P , we get the following linear differential operator of m th-order:

$$\mathbf{T}_P(\mathbf{u}_1, \dots, \mathbf{u}_n) = T_{\beta_1 \dots \beta_n}^{\alpha_1 \dots \alpha_m}(P) u_1^{\beta_1}(P) \dots u_n^{\beta_n}(P) \frac{\partial^m}{\partial x_1^{\alpha_1} \dots \partial x_m^{\alpha_m}}.$$

Here, $\mathbf{u}_j = u_j^\alpha(P) \partial_\alpha$ and $x_k := (x_k^0, x_k^1, x_k^2, x_k^3)$ where $j = 1, \dots, n$. Moreover, $k = 1, \dots, m$. The map

$$P \mapsto \mathbf{T}_P$$

is called a tensor field of type (m, n) on the space-time manifold \mathcal{M}^4 . The elementary proof of statements (K1) through (K3) based on the chain rule will be given in Chap. 8.

Observe the following. By the Einstein convention, we sum over equal upper and lower Greek indices from 0 to 3.

The essential feature is that both the expressions from (K1) and (K3) have no free indices anymore.

The summation kills the indices. Therefore, we summarize (K1) through (K3) under the slogan “principle of killing indices.”

Examples. Taking for granted that $g_{\alpha\beta}, g^{\alpha\beta}, T_{\alpha\beta}$, and $R_{\alpha\beta\gamma}^\delta$ are tensorial families, which will be shown in Chap. 8, it follows from (K2) that the following expressions are tensorial families:

- $R_{\alpha\beta\gamma\delta} = R_{\alpha\beta\gamma}^\sigma g_{\sigma\delta}$,
- $R_{\alpha\delta} = R_{\alpha\beta\gamma\delta} g^{\beta\gamma}$,
- $R = R_{\alpha\beta} g^{\alpha\beta} = R_\alpha^\alpha$ (trace).

In particular, R is an invariant function on the space-time manifold \mathcal{M}^4 , by (K1). In addition, it follows from (K3) that the following expressions are invariantly defined tensor fields on the space-time manifold \mathcal{M}^4 :

- $\mathbf{g} := g_{\alpha\beta} dx^\alpha \otimes dx^\beta$ (metric tensor field),
- $\text{Ric}(\mathbf{g}) := R_{\alpha\beta} dx^\alpha \otimes dx^\beta$ (Ricci tensor field),
- $\mathbf{R} := R_{\alpha\beta\gamma}^\delta dx^\alpha \otimes dx^\beta \otimes dx^\gamma \otimes \partial_\delta$ (Riemann curvature tensor field),
- $\mathcal{R} := R_{\alpha\beta\gamma\delta} dx^\alpha \otimes dx^\beta \otimes dx^\gamma \otimes dx^\delta$ (metric Riemann curvature tensor field).

In what follows, we want to discuss the geometric meaning of both the Riemann curvature tensor field and the metric Riemann curvature tensor field.

Levi-Civita’s parallel transport of velocity vectors. In 1917, Levi-Civita (1873–1941) published a fundamental paper.³¹

³¹ T. Levi-Civita, The notion of parallel transport in manifolds and its geometric consequences for the Riemann curvature, *Rend. Palermo* **42** (1917), 73–205 (in Italian).

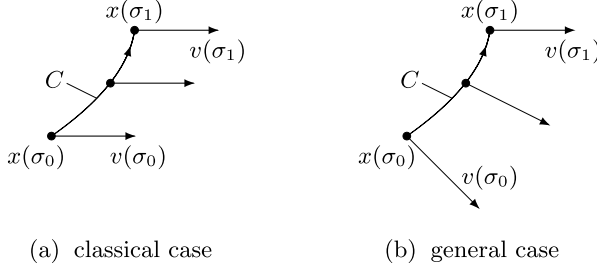


Fig. 0.2. Parallel transport of the vector v along the curve C

*He proved that the Riemann curvature tensor can be obtained by parallel transport of velocity vectors along a sufficiently small closed curve.*³²

This observation was crucial for the development of modern differential geometry and modern physics. Let us discuss this. Consider a curve C on the space-time manifold \mathcal{M}^4 . With respect to a given local coordinate system, the curve C is given by the equation

$$x = x(\sigma), \quad \sigma_0 \leq \sigma \leq \sigma_1.$$

In what follows we will write

$$v = \begin{pmatrix} v^0 \\ v^1 \\ v^2 \\ v^3 \end{pmatrix} \quad \text{and} \quad \mathbf{v} = v^\alpha \partial_\alpha.$$

That is, the components of the velocity vector \mathbf{v} (which lives in the tangent space $T_P\mathcal{M}^4$) are the entries of the column matrix v (which represents a vector in \mathbb{R}^4). If the curve C is a geodesic line, then it satisfies the differential equation

$$\ddot{x}^\gamma(\sigma) = -\dot{x}^\alpha(\sigma) \Gamma_{\alpha\beta}^\gamma(x(\sigma)) \dot{x}^\beta(\sigma), \quad \sigma_0 \leq \sigma \leq \sigma_1, \quad \gamma = 0, 1, 2, 3.$$

Setting $v^\gamma := \dot{x}^\gamma(\sigma)$, we get the differential equation

$$\dot{v}^\gamma(\sigma) = -\dot{x}^\alpha(\sigma) \Gamma_{\alpha\beta}^\gamma(x(\sigma)) \cdot v^\beta(\sigma), \quad \sigma_0 \leq \sigma \leq \sigma_1, \quad \gamma = 0, 1, 2, 3. \quad (0.43)$$

Introducing the real (4×4) -matrix $\mathcal{A}_\alpha := (\Gamma_{\alpha\beta}^\gamma)$, equation (0.43) can be written as

$$\boxed{\dot{v}(\sigma) = -\dot{x}^\alpha(\sigma) \mathcal{A}_\alpha(x(\sigma)) \cdot v(\sigma), \quad \sigma_0 \leq \sigma \leq \sigma_1.} \quad (0.44)$$

This is called the equation of parallel transport. Now to the point. We replace the geodesic line by a general curve C , and we consider general solutions $v = v(\sigma)$ of (0.44). In this setting, we say that the velocity vector $v(\sigma)$ is parallel transported

³² Sophus Lie motivated his approach to differential geometry by the physical picture of the flow of fluid particles on, say, a sphere (e.g., an ocean on earth). In this setting, tangent vectors of the sphere represent velocity vectors of fluid particles. Therefore, in this monograph, tangent vectors of manifolds will be synonymously called velocity vectors.

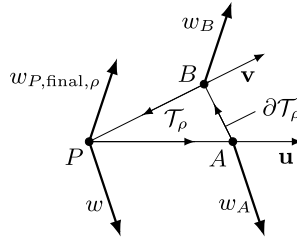


Fig. 0.3. Parallel transport of the vector w along the loop ∂T_ρ

along the curve C . Note that this is a generalization of the classical parallel transport of vectors in the Euclidean manifold (Fig. 0.2). Furthermore, our definition of parallel transport is chosen in such a way that, as a special case, the tangent vectors of a geodesic line C are parallel along this line. The simple geometric meaning of the parallel transport of velocity vectors on a sphere (based on orthogonal projection) will be discussed in Sect. 9.5 on page 593.

Now let us use the notion of parallel transport in order to define the linear operator $\Pi_C : T_P \mathcal{M}^4 \rightarrow T_P \mathcal{M}^4$ given by

$$\boxed{\Pi_C \mathbf{v}(\sigma_0) := \mathbf{v}(\sigma_1)}.$$

More precisely, we proceed as follows (Fig. 0.2(b)):

- We are given the velocity vector $\mathbf{v}(\sigma_0) \in T_P \mathcal{M}^4$. Choosing a fixed local coordinate system, the vector $\mathbf{v}(\sigma_0)$ corresponds to the coordinate matrix $v(\sigma_0)$.
- Solving the differential equation (0.44) of parallel transport, we get the function $v = v(\sigma)$ along the curve $C : x = x(\sigma)$.
- Finally, by definition, the velocity vector $\mathbf{v}(\sigma_1)$ corresponds to $v(\sigma_1)$.

The following fact is crucial. We will prove in Chap. 8 that the transformation law for the Christoffel symbols implies that:

The definition of the operator Π_C does not depend on the choice of the local coordinate system.

In other words, the parallel transport of a velocity vector along a curve C is a geometric property of the space-time manifold \mathcal{M}^4 . In terms of physics, this corresponds to the transport of physical information.

The geometric meaning of the Riemann curvature tensor via parallel transport. Fix the point P of the space-time manifold \mathcal{M}^4 . We are given the two tangent vectors $\mathbf{u}, \mathbf{v} \in T_P \mathcal{M}^4$ at the point P . It is our goal to construct the operator

$$\mathbf{F}_P(\mathbf{u}, \mathbf{v}) : T_P \mathcal{M}^4 \rightarrow T_P \mathcal{M}^4$$

which measures the curvature of \mathcal{M}^4 at the point P with respect to the plane spanned by \mathbf{u} and \mathbf{v} . To this end, choose a fixed, but otherwise arbitrary chart (i.e., a local coordinate system), and consider the situation pictured in Fig. 0.3.

- Fix the scaling factor $\varrho > 0$. Let T_ϱ denote the triangle spanned by the vectors $\varrho \mathbf{u}$ and $\varrho \mathbf{v}$, and let ∂T_ϱ denote the positively oriented boundary curve of T_ϱ .
- Parallel transport of the given vector \mathbf{w} at the initial point P along the closed curve $PABP = \partial T_\varrho$ yields the vector $w_{P,\text{final};\varrho} = \Pi_{\partial T_\varrho} w$ at the final point P .

The following hold.

- Let $\text{meas}(\mathcal{T}_\varrho)$ denote the Euclidean measure of the triangle \mathcal{T}_ϱ . There exists the limit

$$\lim_{\varrho \rightarrow 0} \frac{\Pi_{\partial \mathcal{T}_\varrho} w}{\text{meas}(\mathcal{T}_\varrho)} = w_{P, \text{final}}.$$

- The tangent vector $\mathbf{w}_{P, \text{final}}$ corresponding to the coordinate matrix $w_{P, \text{final}}$ does not depend on the choice of the local coordinates.

We define

$$\boxed{\mathbf{F}_P(\mathbf{u}, \mathbf{v})\mathbf{w} := \mathbf{w}_{P, \text{final}}.}$$

The operator $\mathbf{w} \mapsto \mathbf{F}(\mathbf{u}, \mathbf{v})\mathbf{w}$ is called the Riemann curvature operator.

In physics, we want to use real numbers which can be measured in physical experiments. In order to pass from the Riemann curvature operator to real numbers, it is quite natural to use the (indefinite) inner product on the tangent space $T_P \mathcal{M}^4$ of the space-time manifold at the point P . Explicitly, fix a tangent vector $\mathbf{z} \in T_P \mathcal{M}^4$, and consider the inner product

$$\boxed{\mathcal{R}_P(\mathbf{u}, \mathbf{v}; \mathbf{w}, \mathbf{z}) := \langle \mathbf{F}_P(\mathbf{u}, \mathbf{v})\mathbf{w} | \mathbf{z} \rangle.} \quad (0.45)$$

This is equal to $\mathbf{g}_P(\mathbf{F}_P(\mathbf{u}, \mathbf{v})\mathbf{w}, \mathbf{z})$. The map

$$(\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}) \mapsto \mathcal{R}_P(\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z})$$

is a 4-linear map of the form

$$\mathcal{R}_P : T_P \mathcal{M}^4 \times T_P \mathcal{M}^4 \times T_P \mathcal{M}^4 \times T_P \mathcal{M}^4 \rightarrow \mathbb{R}.$$

This map is called the metric Riemann curvature tensor.

Let $\mathbf{u} = u^\alpha \partial_\alpha$ (together with similar representations of \mathbf{v}, \mathbf{w} , and \mathbf{z}). With respect to local coordinates, we get the following symmetric formulas:

- $\mathbf{F}(\mathbf{u}, \mathbf{v})\mathbf{w} = (R_{\alpha\beta\gamma}^\kappa u^\alpha v^\beta w^\gamma) \partial_\kappa$,
- $\mathcal{R}_P(\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}) = R_{\alpha\beta\gamma\delta} u^\alpha v^\beta w^\gamma z^\delta$,
- $\text{Ric}(\mathbf{u}, \mathbf{z}) := \langle \mathbf{F}(\mathbf{u}, \partial_\beta) \partial_\gamma | \mathbf{z} \rangle g^{\beta\gamma} = R_{\alpha\beta} u^\alpha v^\beta$ (averaging).

Here, the real coefficients are given by

- $R_{\alpha\beta\gamma}^\delta := dx^\delta(\mathbf{F}(\partial_\alpha, \partial_\beta)\partial_\gamma)$,
- $R_{\alpha\beta\gamma\delta} := \mathcal{R}_P(\partial_\alpha, \partial_\beta, \partial_\gamma, \partial_\delta)$,
- $R_{\alpha\delta} := R_{\alpha\beta\gamma\delta} g^{\beta\gamma}$.

Symmetries of the Riemann curvature tensor. We have:

- $\mathbf{F}(\mathbf{u}, \mathbf{v})\mathbf{w} = -\mathbf{F}(\mathbf{v}, \mathbf{u})\mathbf{w}$,
- $\mathbf{F}(\mathbf{u}, \mathbf{v})\mathbf{w} + \mathbf{F}(\mathbf{v}, \mathbf{w})\mathbf{u} + \mathbf{F}(\mathbf{w}, \mathbf{u})\mathbf{v} = 0$ (cyclic permutation),
- $\mathcal{R}_P(\mathbf{u}, \mathbf{v}; \mathbf{w}, \mathbf{z}) = -\mathcal{R}_P(\mathbf{v}, \mathbf{u}; \mathbf{w}, \mathbf{z})$,
- $\mathcal{R}_P(\mathbf{u}, \mathbf{v}; \mathbf{w}, \mathbf{z}) = -\mathcal{R}_P(\mathbf{u}, \mathbf{v}; \mathbf{z}, \mathbf{w})$,
- $\mathcal{R}_P(\mathbf{u}, \mathbf{v}; \mathbf{w}, \mathbf{z}) = \mathcal{R}_P(\mathbf{w}, \mathbf{z}; \mathbf{u}, \mathbf{v})$,
- $\text{Ric}(\mathbf{g})(\mathbf{u}, \mathbf{v}) = \text{Ric}(\mathbf{g})(\mathbf{v}, \mathbf{u})$.

For the coefficients $R_{\alpha\beta\gamma\delta}$ and $R_{\alpha\beta}$, this yields the symmetry relations summarized on page 14.

Riemann's sectional curvature and the geometric meaning of the Riemann curvature tensor. Let $\mathbf{u} = u^\alpha \partial_\alpha$ and $\mathbf{v} = v^\beta \partial_\beta$ be linearly independent tangent vectors of the space-time manifold \mathcal{M}^4 at the point P . The sectional curvature of \mathcal{M}^4 at the point P is defined by ³³

³³ Naturally enough, we assume that $v_P(\mathbf{u}, \mathbf{v}) \neq 0$.

$$K_P(\mathbf{u}, \mathbf{v}) := \frac{\mathcal{R}_P(\mathbf{u}, \mathbf{v}, \mathbf{v}, \mathbf{u})}{v_P(\mathbf{u}, \mathbf{v})^2} \quad (0.46)$$

where $v_P(\mathbf{u}, \mathbf{v})^2 := \mathbf{g}_P(\mathbf{u}, \mathbf{u})\mathbf{g}_P(\mathbf{v}, \mathbf{v}) - \mathbf{g}_P(\mathbf{u}, \mathbf{v})^2$. If \mathbf{u} and \mathbf{v} are space-like vectors, then $v_P(\mathbf{u}, \mathbf{v})$ is the surface area of the parallelogram spanned by the two vectors \mathbf{u} and \mathbf{v} at the point P . It turns out that this sectional curvature only depends on the 2-dimensional plane \mathcal{P} spanned by the tangent vectors \mathbf{u} and \mathbf{v} .

- If two 2-dimensional submanifolds \mathcal{S} and \mathcal{S}' of the space-time manifold \mathcal{M}^4 have the same tangent plane at the point P , then they possess the same sectional curvature at the point P .
- If the plane \mathcal{P} is space-like, then the sectional curvature coincides with the Gaussian curvature of \mathcal{S} and \mathcal{S}' at the point P .

This sectional curvature was introduced by Riemann in his seminal lecture “On the hypotheses which lie at the foundation of geometry” in 1854.³⁴

It was the idea of Riemann to describe the curvature of a higher-dimensional manifold \mathcal{M} at the point P by studying the Gaussian curvature of all possible two-dimensional submanifolds \mathcal{M}^2 of \mathcal{M} at the point P .

Let $\mathcal{K}(\mathbf{u}, \mathbf{v})_P := v_P(\mathbf{u}, \mathbf{v})^2 K_P(\mathbf{u}, \mathbf{v})$. Then

$$\begin{aligned} \mathcal{R}(\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}) &= \mathcal{K}(\mathbf{u} + \mathbf{z}, \mathbf{v} + \mathbf{w}) - \mathcal{K}(\mathbf{u} + \mathbf{z}, \mathbf{v}) - \mathcal{K}(\mathbf{u} + \mathbf{z}, \mathbf{w}) \\ &\quad - \mathcal{K}(\mathbf{u}, \mathbf{v} + \mathbf{w}) - \mathcal{K}(\mathbf{z}, \mathbf{v} + \mathbf{w}) + \mathcal{K}(\mathbf{u}, \mathbf{w}) + \mathcal{K}(\mathbf{z}, \mathbf{v}) \\ &\quad - \mathcal{K}(\mathbf{v} + \mathbf{z}, \mathbf{u} + \mathbf{w}) + \mathcal{K}(\mathbf{v} + \mathbf{z}, \mathbf{u}) + \mathcal{K}(\mathbf{v} + \mathbf{z}, \mathbf{w}) \\ &\quad - \mathcal{K}(\mathbf{v}, \mathbf{u} + \mathbf{w}) + \mathcal{K}(\mathbf{z}, \mathbf{u} + \mathbf{w}) - \mathcal{K}(\mathbf{v}, \mathbf{w}) - \mathcal{K}(\mathbf{z}, \mathbf{u}). \end{aligned}$$

This key formula tells us that the sectional curvature determines the Riemann curvature tensor which describes the whole curvature.

Two Fundamental Approaches to Differential Geometry

There exist the following two different approaches to differential geometry, namely,

- (I) Gauss’ approach by means of symmetric tensors, and
- (II) Élie Cartan’s approach by means of antisymmetric tensors (also called differential forms).

The Einstein equation $\text{Ric}(\mathbf{g}) = \kappa_G(\mathbf{T} - \frac{1}{2} \text{tr}(\mathbf{T}) \cdot \mathbf{g})$ (for the motion of the gravitational field by means of the symmetric Ricci tensor $\text{Ric}(\mathbf{g})$) is formulated in the spirit of Gauss (see page 23). In what follows, we will study Cartan’s approach based on the structural equation and its integrability condition (the Bianchi identity). We will sketch the following ideas:

- Yang’s matrix trick,
- Cartan’s local structural equation, and
- Cartan’s global structural equation.

Modern differential geometry is based on Cartan’s approach. The decisive advantage of Cartan’s approach is that it allows the use of symmetry groups (also called gauge groups) in a very flexible way.

³⁴ B. Riemann, Über die Hypothesen, welche der Geometrie zugrundeliegen, Göttinger Abhandlungen **13** (1854), 272–287 (in German). An English translation can be found in M. Spivak, A Comprehensive Introduction to Differential Geometry, Vol. 2, Publish or Perish, Boston.

Yang's Matrix Trick

While preparing a lecture on Einstein's general relativity theory in the 1960s, Yang discovered that the fundamental equations for the components of the Riemann curvature tensor,

$$R_{\alpha\beta\gamma}^{\delta} := \partial_{\alpha}\Gamma_{\beta\gamma}^{\delta} - \partial_{\beta}\Gamma_{\alpha\gamma}^{\delta} + \Gamma_{\alpha\mu}^{\delta}\Gamma_{\beta\gamma}^{\mu} - \Gamma_{\beta\mu}^{\delta}\Gamma_{\alpha\gamma}^{\mu}, \quad (0.47)$$

coincide with the Yang–Mills field equations

$$\mathcal{F}_{\alpha\beta} := \partial_{\alpha}\mathcal{A}_{\beta} - \partial_{\beta}\mathcal{A}_{\alpha} + [\mathcal{A}_{\alpha}, \mathcal{A}_{\beta}] \quad (0.48)$$

with the Lie product $[\mathcal{A}_{\alpha}, \mathcal{A}_{\beta}] := \mathcal{A}_{\alpha}\mathcal{A}_{\beta} - \mathcal{A}_{\beta}\mathcal{A}_{\alpha}$. Here, the indices $\alpha, \beta, \gamma, \delta$ run from 0 to 3. Observe that formula (0.47) can be regarded as a generalization of Gauss' theorema egregium to higher dimensions. In order to obtain Yang's result, let us introduce the following matrices:

$$\mathcal{A}_{\alpha} := (\Gamma_{\alpha\gamma}^{\delta}), \quad \mathcal{F}_{\alpha\beta} := (R_{\alpha\beta\gamma}^{\delta})$$

where the upper index δ numbers the rows, and the lower index γ numbers the columns. Explicitly, this means the following:

(i) Christoffel matrices (connection matrices):

$$\mathcal{A}_{\alpha} := \begin{pmatrix} \Gamma_{\alpha 0}^0 & \Gamma_{\alpha 1}^0 & \Gamma_{\alpha 2}^0 & \Gamma_{\alpha 3}^0 \\ \Gamma_{\alpha 0}^1 & \Gamma_{\alpha 1}^1 & \Gamma_{\alpha 2}^1 & \Gamma_{\alpha 3}^1 \\ \Gamma_{\alpha 0}^2 & \Gamma_{\alpha 1}^2 & \Gamma_{\alpha 2}^2 & \Gamma_{\alpha 3}^2 \\ \Gamma_{\alpha 0}^3 & \Gamma_{\alpha 1}^3 & \Gamma_{\alpha 2}^3 & \Gamma_{\alpha 3}^3 \end{pmatrix}.$$

(ii) Riemann curvature matrices:

$$\mathcal{F}_{\alpha\beta} := \begin{pmatrix} R_{\alpha\beta 0}^0 & R_{\alpha\beta 1}^0 & R_{\alpha\beta 2}^0 & R_{\alpha\beta 3}^0 \\ R_{\alpha\beta 0}^1 & R_{\alpha\beta 1}^1 & R_{\alpha\beta 2}^1 & R_{\alpha\beta 3}^1 \\ R_{\alpha\beta 0}^2 & R_{\alpha\beta 1}^2 & R_{\alpha\beta 2}^2 & R_{\alpha\beta 3}^2 \\ R_{\alpha\beta 0}^3 & R_{\alpha\beta 1}^3 & R_{\alpha\beta 2}^3 & R_{\alpha\beta 3}^3 \end{pmatrix}.$$

Using the multiplication of matrices, it follows immediately that the Yang–Mills field equation (0.48) corresponds to the Riemann equation (0.47).

Cartan's Local Structural Equation

In order to get insight in differential geometry, use differential forms and employ their invariance properties.

Folklore

In order to kill indices, let us define the following differential forms:

- $\mathcal{A} := \mathcal{A}_{\alpha} dx^{\alpha}$ (local connection form on the space-time manifold \mathcal{M}^4),
- $\mathcal{F} := \frac{1}{2}\mathcal{F}_{\alpha\beta} dx^{\alpha} \wedge dx^{\beta}$ (local curvature form on \mathcal{M}^4).

Here, $\mathcal{F}_{\alpha\beta} = -\mathcal{F}_{\beta\alpha}$ for all indices $\alpha, \beta = 0, 1, 2, 3$. Note that \mathcal{A} and \mathcal{F} are differential forms with the (4×4) -matrices \mathcal{A}_{α} and $\mathcal{F}_{\alpha\beta}$ as coefficients. This can also be written as

$$\mathcal{A} = (\omega_{\gamma}^{\delta}), \quad \mathcal{F} = (\Omega_{\gamma}^{\delta})$$

with the differential forms

- $\omega_\gamma^\delta := \Gamma_{\alpha\gamma}^\delta dx^\alpha$ (Cartan's local connection forms), and
- $\Omega_\gamma^\delta := \frac{1}{2} R_{\alpha\beta\gamma}^\delta dx^\alpha \wedge dx^\beta$ (Cartan's local curvature forms).

Calculus for matrices with differential forms as entries. If a, b, c, d and e, f, g, h are complex numbers, then we have the following classical matrix product:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix}.$$

Now suppose that all the entries a, b, \dots are differential forms. Then the wedge products $a \wedge e, \dots$ of entries are well defined. This motivates the following definition of the wedge product of matrices with differential forms as entries:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \wedge \begin{pmatrix} e & f \\ g & h \end{pmatrix} := \begin{pmatrix} a \wedge e + b \wedge g & a \wedge f + b \wedge h \\ c \wedge e + d \wedge g & c \wedge f + d \wedge h \end{pmatrix}. \quad (0.49)$$

That is, we merely replace the classical product of entries by the wedge product of entries.

Cartan's local structural equation. We claim that the differential forms \mathcal{A} and \mathcal{F} satisfy the following two elegant equations:

(C) Cartan's local structural equation:

$$\boxed{\mathcal{F} = d\mathcal{A} + \mathcal{A} \wedge \mathcal{A}.} \quad (0.50)$$

(B) Bianchi's local identity (integrability condition to (C)):

$$\boxed{d\mathcal{F} = \mathcal{F} \wedge \mathcal{A} - \mathcal{A} \wedge \mathcal{F}.} \quad (0.51)$$

Proof. Ad (C). This is nothing else than a clever reformulation of the classical curvature relation

$$R_{\alpha\beta\gamma}^\delta := \partial_\alpha \Gamma_{\beta\gamma}^\delta - \partial_\beta \Gamma_{\alpha\gamma}^\delta + \Gamma_{\alpha\mu}^\delta \Gamma_{\beta\gamma}^\mu - \Gamma_{\beta\mu}^\delta \Gamma_{\alpha\gamma}^\mu$$

in terms of Cartan's calculus of differential forms invented by Cartan in 1899.³⁵ In fact, it follows from $\omega_\gamma^\delta = \Gamma_{\beta\gamma}^\delta dx^\beta$ and the antisymmetry of the Grassmann product (also called the wedge product) $dx^\alpha \wedge dx^\beta = -dx^\beta \wedge dx^\alpha$ that

$$\begin{aligned} d\omega_\gamma^\delta &= d\Gamma_{\beta\gamma}^\delta \wedge dx^\beta = \partial_\alpha \Gamma_{\beta\gamma}^\delta dx^\alpha \wedge dx^\beta \\ &= \frac{1}{2} (\partial_\alpha \Gamma_{\beta\gamma}^\delta - \partial_\beta \Gamma_{\alpha\gamma}^\delta) dx^\alpha \wedge dx^\beta \end{aligned}$$

and

$$\begin{aligned} \omega_\mu^\delta \wedge \omega_\gamma^\mu &= \Gamma_{\alpha\mu}^\delta \Gamma_{\beta\gamma}^\mu dx^\alpha \wedge dx^\beta \\ &= \frac{1}{2} (\Gamma_{\alpha\mu}^\delta \Gamma_{\beta\gamma}^\mu - \Gamma_{\beta\mu}^\delta \Gamma_{\alpha\gamma}^\mu) dx^\alpha \wedge dx^\beta. \end{aligned}$$

Using $\Omega_\gamma^\delta = \frac{1}{2} R_{\alpha\beta\gamma}^\delta dx^\alpha \wedge dx^\beta$, we obtain Cartan's system of structural equations

$$\boxed{\Omega_\gamma^\delta = d\omega_\gamma^\delta + \omega_\mu^\delta \wedge \omega_\gamma^\mu, \quad \gamma, \delta = 0, 1, 2, 3.} \quad (0.52)$$

³⁵ Grassmann (1809–1877), Élie Cartan (1869–1951).

In order to kill the indices, we introduce the matrices $\mathcal{A} = (\omega_\gamma^\delta)$ and $\mathcal{F} = (\Omega_\gamma^\delta)$. Then Cartan's system (0.52) can be elegantly written as

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

Ad (B). We will use the following properties of the calculus for differential forms (with real coefficients). Let ω , ϱ and τ be differential forms of degree p , r , and s , respectively. Then, the wedge product has the following properties:

- $(\omega \wedge \varrho) \wedge \tau = \omega \wedge (\varrho \wedge \tau)$ (associativity),
- $\omega \wedge \varrho = (-1)^{pr} \varrho \wedge \omega$ (graded anticommutativity),
- $d\omega = 0$ (the Poincaré cohomology rule),
- $d(\omega \wedge \varrho) = d\omega \wedge \varrho + (-1)^p \omega \wedge d\varrho$ (the graded Leibniz product rule).

These properties induce the corresponding rules for the wedge product of matrices. In particular, applying the Cartan differential “ d ” to (C), we get

$$d\mathcal{F} = d(d\mathcal{A}) + d\mathcal{A} \wedge \mathcal{A} - \mathcal{A} \wedge d\mathcal{A}.$$

By the Poincaré cohomology rule, $d(d\mathcal{A}) = 0$. Since $\mathcal{F} = d\mathcal{A} + \mathcal{A} \wedge \mathcal{A}$, we get

$$d\mathcal{F} = (\mathcal{F} - \mathcal{A} \wedge \mathcal{A}) \wedge \mathcal{A} - \mathcal{A} \wedge (\mathcal{F} - \mathcal{A} \wedge \mathcal{A}) = \mathcal{F} \wedge \mathcal{A} - \mathcal{A} \wedge \mathcal{F}.$$

□

Gauge transformations. Let us consider the change

$$x^{\alpha'} = x^{\alpha'}(x^0, x^1, x^2, x^3), \quad \alpha = 0, 1, 2, 3$$

of local coordinates. We want to determine the transformation laws for the family of Christoffel symbols $\Gamma_{\alpha\lambda}^\kappa$ and the family of Riemann symbols $R_{\lambda\alpha\beta}^\kappa$. We will use the notation introduced on page 26. Let us first describe an elementary brute-force method based on completely elementary, but lengthy computations based on the chain rule in classical differential calculus.

- We start with the transformation law³⁶

$$g_{\alpha'\beta'} = G_{\alpha'}^\alpha G_{\beta'}^\beta g_{\alpha\beta}$$

for the components $g_{\alpha\beta}$ of the metric tensor. Using matrices, this means that

$$(g_{\alpha'\beta'}) = G^d (g_{\alpha\beta}) G.$$

- For the inverse matrix, we get

$$(g^{\alpha'\beta'}) = (g_{\alpha'\beta'})^{-1} = G^{-1} (g_{\alpha\beta})^{-1} (G^d)^{-1} = G^{-1} (g^{\alpha\beta}) (G^{-1})^d.$$

This implies

$$g^{\alpha'\beta'} = G_{\alpha'}^{\alpha'} G_{\beta'}^{\beta'} \cdot g^{\alpha\beta}.$$

Thus, $g^{\alpha\beta}$ is a tensorial family of type $(2, 0)$.

- We use $\Gamma_{\alpha\beta}^\delta := \frac{1}{2}(\partial_\alpha g_{\beta\sigma} + \partial_\beta g_{\alpha\sigma} - \partial_\sigma g_{\alpha\beta}) g^{\sigma\kappa}$ in order to get the transformation formula for the Christoffel symbols:

$$\Gamma_{\alpha'\beta'}^{\delta'} = G_{\alpha'}^{\delta'} G_{\beta'}^{\delta'} \cdot \Gamma_{\alpha\beta}^\delta - G_{\alpha'}^{\alpha'} G_{\beta'}^{\beta'} (\partial_\alpha G_{\beta'}^{\delta'}). \quad (0.53)$$

³⁶ More precisely, $g_{\alpha'\beta'}(x') = G_{\alpha'}^\alpha(x) G_{\beta'}^\beta(x) g_{\alpha\beta}(x)$.

- From (0.47), we obtain the following transformation formula for the Riemann symbols:

$$R_{\alpha'\beta'\gamma'}^{\delta'} = G_{\delta}^{\delta'} G_{\alpha'}^{\alpha} G_{\beta'}^{\beta} G_{\gamma'}^{\gamma} \cdot R_{\alpha\beta\gamma}^{\delta}. \quad (0.54)$$

Relation (0.47) tells us that $R_{\alpha\beta\gamma}^{\delta}$ forms a tensorial family of type (1, 3). Moreover, it follows from (0.53) that the Christoffel family $\Gamma_{\alpha\beta}^{\delta}$ is not a tensorial family.³⁷ In contrast to this brute-force method, a much simpler proof of (0.53) and (0.47) will be given in Chap. 8 based on the inverse index principle (see page 505). More elegantly, using the language of matrices, the transformation formulas (0.53) and (0.47) can be written in the following way:

- (i) $\dot{x}' = G\dot{x}$ (transformation law for the velocity components $\dot{x}^{\alpha}(\sigma)$),
- (ii) $\mathcal{A}' = G\mathcal{A}G^{-1} - (dG)G^{-1}$ (transformation law for the connection form), and
- (iii) $\mathcal{F}' = G\mathcal{F}G^{-1}$ (transformation law for the curvature form).³⁸

Here, we set $\mathcal{A}(x) := \mathcal{A}_{\alpha}(x)dx^{\alpha} = (\Gamma_{\alpha\gamma}^{\delta}(x)dx^{\alpha})$ and

$$\mathcal{A}'(x') := (\Gamma_{\alpha'\gamma'}^{\delta'}(x') dx^{\alpha'}), \quad x' = x'(x).$$

Note that $\mathcal{A}(x)$ is a matrix with the differential forms $\Gamma_{\alpha\gamma}^{\delta}(x)dx^{\alpha}$ as entries; the upper index δ numbers the rows, and the lower index γ numbers the columns. Furthermore,

$$\mathcal{F} := \frac{1}{2}\mathcal{F}_{\alpha\beta} dx^{\alpha} \wedge dx^{\beta} = \left(\frac{1}{2}R_{\alpha\beta\gamma}^{\delta} dx^{\alpha} \wedge dx^{\beta}\right),$$

and

$$\mathcal{F}'(x') := \left(\frac{1}{2}R_{\alpha'\beta'\gamma'}^{\delta'}(x') dx^{\alpha'} \wedge dx^{\beta'}\right).$$

Finally, recall that $G = (G_{\delta}^{\delta'})$ and $G^{-1} = (G_{\delta'}^{\delta})$. The proof of (ii), (iii) above will be given in Chap. 8.

Cartan's Global Structural Equation

Extend the space-time manifold \mathcal{M}^4 to its frame bundle.
Folklore

In order to arrive at a global approach, we pass from the local structural equation $\mathcal{F} = d\mathcal{A} + \mathcal{A} \wedge \mathcal{A}$ considered in (0.50) to the global structural equation

$$\boxed{\mathcal{F} = d\mathcal{A}}. \quad (0.55)$$

Moreover, we add the global integrability condition

$$\boxed{d\mathcal{F} = 0} \quad (0.56)$$

which is called the global Bianchi identity. Let us discuss this.

Extension of the space-time manifold \mathcal{M}^4 to the frame bundle $F(\mathcal{M}^4)$.
The two differential forms

³⁷ The Christoffel family is also called the connection family.

³⁸ More precisely, $x'(\sigma) = G(x(\sigma))x(\sigma)$,

$$\mathcal{A}'(x') = (G\mathcal{A}G^{-1} - dG \cdot G^{-1})(x),$$

and $\mathcal{F}'(x') = (G\mathcal{F}G^{-1})(x)$. Note that the prime refers to the transformed coordinates $x' = (x^{1'}, \dots, x^{n'})$, but it does not refer to any derivative.

- A (global connection form), and
- F (global curvature form)

are not defined on the space-time manifold, but on the frame bundle. By definition, the frame bundle $F(\mathcal{M}^4)$ of the space-time manifold \mathcal{M}^4 consists of all the tuples

$$(P, \mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) \quad (0.57)$$

where P is an arbitrary point of \mathcal{M}^4 , and $\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ is an arbitrary basis of the tangent space $T_P\mathcal{M}^4$ of \mathcal{M}^4 at the point P . Choose local coordinates (x^0, x^1, x^2, x^3) which live in the open set U of \mathbb{R}^4 . Since $\partial_0, \partial_1, \partial_2, \partial_3$ form a basis of the tangent space $T_P\mathcal{M}^4$, there exist real numbers $G_\alpha^\beta(x)$ depending on x such that

$$\mathbf{b}_\alpha = G_\alpha^\beta(x) \partial_\beta, \quad \alpha = 0, 1, 2, 3.$$

Introducing the matrix $G(x) := (G_\alpha^\beta(x))$, we get the matrix equation

$$(\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) = (\partial_0, \partial_1, \partial_2, \partial_3) G(x)$$

where $G(x)$ is an invertible real (4×4) -matrix, that is, $G(x) \in GL(4, \mathbb{R})$. The tuple $(x, G(x))$ is called the local coordinate of the point (0.57) of the frame bundle $F(\mathcal{M}^4)$. Obviously,

$$(x, G(x)) \in U \times GL(4, \mathbb{R}).$$

Parallel transport on the space-time manifold in terms of the frame bundle. We are given the curve

$$C : P = P(\sigma), \quad \sigma_0 \leq \sigma \leq \sigma_1,$$

on the space-time manifold \mathcal{M}^4 . With respect to local coordinates, this curve corresponds to the map $\sigma \mapsto x(\sigma)$. Consider the differential equation

$$\dot{G}(\sigma) = -\dot{x}^\alpha(\sigma) \mathcal{A}_\alpha(x(\sigma)) \cdot G(\sigma), \quad \sigma_0 \leq \sigma \leq \sigma_1, \quad (0.58)$$

with the initial condition $G(\sigma_0) = I$. Let $G = G(\sigma)$ be the unique solution of the initial-value problem for (0.58). We are given the tangent vector $\mathbf{v}_0 = v_0^\alpha \partial_\alpha$ at the point $P(\sigma_0)$. We set $v(\sigma) := G(\sigma)v_0$. Then the differential equation of parallel transport

$$\dot{v}(\sigma) = -\dot{x}^\alpha(\sigma) \mathcal{A}_\alpha(x(\sigma)) \cdot v(\sigma), \quad \sigma_0 \leq \sigma \leq \sigma_1,$$

is satisfied. Consequently, setting

$$\mathbf{v}(\sigma) := v^\alpha(\sigma) \partial_\alpha,$$

we obtain the tangent vector $\mathbf{v}(\sigma)$ at the curve point $P(\sigma)$, and this family represents a parallel transport of velocity vectors along the curve C . We will show later on that this parallel transport can be used to define quite naturally the differential form A, the covariant differential DA and hence the curvature form F on the frame bundle.

Restriction of the global curvature form F to the local curvature form \mathcal{F} and gauge transformations. We want to discuss briefly the relation between the global differential forms A and F on the frame bundle $F(\mathcal{M}^4)$ and the (local) differential forms \mathcal{A} and \mathcal{F} on the space time-manifold \mathcal{M}^4 . Setting

$$s(x) := (x, I),$$

we get the map $s : U \rightarrow U \times GL(4, \mathbb{R})$. The corresponding pull-backs of the differential forms yield

- $s^*A = \mathcal{A}$, and
- $s^*F = \mathcal{F}$.

Moreover, for given function $x \mapsto G(x)$, we set

$$s(x) := (x, G(x)).$$

This corresponds to the choice of frames depending on the point P related to the local coordinate x . Then we get:

- $s^*A(x) = G(x)\mathcal{A}(x)G(x)^{-1} - dG(x) \cdot G(x)^{-1}$, and
- $s^*F(x) = G(x)\mathcal{F}(x)G(x)^{-1}$.

This yields precisely the gauge transformation formulas for \mathcal{A} and \mathcal{F} summarized on page 38, namely,

- $\mathcal{A}'(x) = G(x)\mathcal{A}(x)G(x)^{-1} - dG(x) \cdot G(x)^{-1}$, and
- $\mathcal{F}'(x) = G(x)\mathcal{F}(x)G(x)^{-1}$.

Élie Cartan's method of moving frames. The procedure sketched above is called Cartan's method of moving frames. In terms of mathematics, Cartan's global structural equation (0.55) represents a generalization of the basic formula (0.47) which relates the Riemann curvature tensor to the Christoffel symbols and their first partial derivatives.

In terms of physics, the global curvature form F corresponds to the gravitational force. The structural equation (0.55) relates the gravitational force F (i.e., the global curvature form) to the so-called potential A (i.e., the global connection form). In physics, the use of the frame bundle in general relativity is called the tetrad formalism.

The general form of Cartan's approach (based on frame bundles) allows us to transform the local coordinates of the space-time manifold \mathcal{M}^4 and the local coordinates of the frames in a separate way. This yields optimal flexibility.

Covariant Partial Derivative and the Classical Ricci Calculus

Replace the classical derivatives $\ddot{x}^\alpha(\sigma)$ and $\partial_\alpha v^\beta$ by the covariant derivatives $\frac{D\dot{x}(\sigma)}{d\sigma}$ and $\nabla_\alpha v^\beta$, respectively.

In contrast to the classical partial derivative ∂_α , the covariant partial derivative ∇_α has the very useful property that it sends tensorial families again to tensorial families.

Folklore

Classical identities for partial derivatives. Let $\Theta : U \rightarrow \mathbb{R}$ and

$$u^\kappa, v^\kappa, w^\kappa : U \rightarrow \mathbb{R}^4, \quad \kappa = 0, 1, 2, 3$$

be smooth functions where U is a nonempty open subset of \mathbb{R}^4 (e.g., $U = \mathbb{R}^4$). Choose the coordinates x^0, x^1, x^2, x^3 on \mathbb{R}^4 , and recall the notation $\partial_\alpha := \frac{\partial}{\partial x^\alpha}$ for the partial derivative with respect to the variable x^α . Set

$$d_{v(x)}\Theta(x) := v^\alpha(x)\partial_\alpha\Theta(x).$$

This is called the directional derivative of the function Θ at the point x with respect to the direction $v(x)$. Let $x = x(\sigma)$ be a smooth curve with $x^\alpha(0) = x_0^\alpha$ and $\dot{x}^\alpha(0) = v^\alpha$ for all $\alpha = 0, 1, 2, 3$. Then, by the chain rule,

$$d_v \Theta(x_0) = \frac{d\Theta(x(\sigma))}{d\sigma} \Big|_{\sigma=0}.$$

This motivates the designation “directional derivative.” Furthermore, set

$$v = \begin{pmatrix} v^0 \\ v^1 \\ v^2 \\ v^3 \end{pmatrix}, \quad w = \begin{pmatrix} w^0 \\ w^1 \\ w^2 \\ w^3 \end{pmatrix}, \quad d_{v(x)} w(x) := \begin{pmatrix} v^\alpha(x) \partial_\alpha w^0(x) \\ v^\alpha(x) \partial_\alpha w^1(x) \\ v^\alpha(x) \partial_\alpha w^2(x) \\ v^\alpha(x) \partial_\alpha w^3(x) \end{pmatrix}.$$

Here, $d_{v(x)} w(x)$ is called the directional derivative of the function $w : U \rightarrow \mathbb{R}^4$ at the point x with respect to the direction $v(x)$. Finally, introduce the Lie product

$$\boxed{[v, w] := d_v w - d_w v.} \quad (0.59)$$

In addition, we introduce the symbol $[\partial_\alpha, \partial_\beta]_-$ in the sense of linear operators. That is, we set $[\partial_\alpha, \partial_\beta]_- v := \partial_\alpha \partial_\beta v - \partial_\beta \partial_\alpha v = 0$. Then we have the following trivial identities, which will be generalized to nontrivial identities later on:

(I) Trivial Lie product: $[\partial_\alpha, \partial_\beta]_- = 0$.

(II) Trivial Jacobi identity:

$$([\partial_\alpha, [\partial_\beta, \partial_\gamma]_-]_- + ([\partial_\beta, [\partial_\gamma, \partial_\alpha]_-]_- + ([\partial_\gamma, [\partial_\alpha, \partial_\beta]_-]_-)v = 0. \quad (0.60)$$

(III) Trivial Bianchi identity:

$$(\partial_\alpha [\partial_\beta, \partial_\gamma]_- + \partial_\beta [\partial_\gamma, \partial_\alpha]_- + \partial_\gamma [\partial_\alpha, \partial_\beta]_-)v = 0. \quad (0.61)$$

(IV) The key Lie relation:³⁹

$$\boxed{d_u(d_v w) - d_v(d_u w) = d_{[u, v]} w.} \quad (0.62)$$

(V) The Lie algebra $C^\infty(U, \mathbb{R}^4)$: The real linear space $C^\infty(U, \mathbb{R}^4)$ of all smooth functions

$$v : U \rightarrow \mathbb{R}^4$$

forms a real Lie algebra with respect to the Lie product (0.59). Explicitly, this means that $C^\infty(U, \mathbb{R}^4)$ is a real linear space. Furthermore, for all functions $u, v, w \in C^\infty(U, \mathbb{R}^4)$ and all real numbers λ, μ , the following hold:

- $[u, v] \in C^\infty(U, \mathbb{R}^4)$ (consistency),
- $[\lambda u + \mu v, w] = \lambda[u, w] + \mu[v, w]$ (distributivity),
- $[v, w] = -[w, v]$ (antisymmetry),
- $[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0$ (Jacobi identity).

(VI) The Lie algebra $D(C^\infty(U, \mathbb{R}^4))$: By definition, the symbol $D(C^\infty(U, \mathbb{R}^4))$ denotes the set of all linear differential operators

$$v^\alpha(x) \partial_\alpha, \quad x \in U$$

with smooth coefficient functions $v^\alpha : U \rightarrow \mathbb{R}$. This coincides with the set of all differential operators

³⁹ Leibniz (1646–1717), Jacobi (1804–1851), Lie (1842–1899), Bianchi (1856–1928).

$$d_v : C^\infty(U, \mathbb{R}^4) \rightarrow C^\infty(U, \mathbb{R}^4)$$

where $v \in C^\infty(U, \mathbb{R}^4)$. The real linear space $D(C^\infty(U, \mathbb{R}^4))$ becomes a real Lie algebra with respect to the Lie product $[d_u, d_v]_-$ in the sense of operators. That is,

$$[d_u, d_v]_- w = d_u(d_v w) - d_v(d_u w) \quad (0.63)$$

for all $w \in C^\infty(U, \mathbb{R}^4)$. Observe the following crucial fact for the Lie theory of partial differential equations. By (IV),

$$[d_u, d_v]_- w = d_{[u, v]} w.$$

Consequently, $[d_u, d_v]_-$ is not a second-order differential operator, but only a first-order partial differential operator because of the cancellation of partial derivatives of second order. Therefore, $[d_u, d_v]_- w$ is an element of the linear space $D(C^\infty(U, \mathbb{R}^4))$. The Jacobi identity

$$[A, [B, C]_-]_- w + [B, [C, A]_-]_- w + [C, [A, B]_-]_- w = 0$$

on $D(C^\infty(U, \mathbb{R}^4))$ follows from (0.63). This is a special case of the general fact that the Jacobi identity is always satisfied for linear operators on linear spaces (see (0.66)).

(VII) Leibniz rule: For all $v \in C^\infty(U, \mathbb{R}^4)$ and all smooth functions $\Theta : U \rightarrow \mathbb{R}$, we have the product rule:

$$d_v(\Theta w) = (d_v \Theta)w + \Theta d_v w. \quad (0.64)$$

(VIII) Differential: Set $dw := \partial_\alpha w dx^\alpha$. Then

$$d_v w = (dw)(v).$$

In fact, $dx^\alpha(v) = v^\alpha$. Hence $dw(v) = v^\alpha \partial_\alpha w$.

Proof. All the statements follow by using elementary computations based on the key relation

$$\boxed{\partial_\alpha \partial_\beta = \partial_\beta \partial_\alpha, \quad \alpha, \beta = 0, 1, 2, 3.} \quad (0.65)$$

This is the mnemonic formulation of the commutativity property

$$\frac{\partial}{\partial x^\alpha} \left(\frac{\partial \Theta}{\partial x^\beta} \right) = \frac{\partial}{\partial x^\beta} \left(\frac{\partial \Theta}{\partial x^\alpha} \right)$$

for the partial derivatives of smooth real-valued functions Θ .

For example, let us prove the key Lie relation (IV). Observe first that

$$d_u(d_v w) = u^\alpha \partial_\alpha (v^\beta \partial_\beta w) = u^\alpha v^\beta \partial_\alpha \partial_\beta w + u^\alpha \partial_\alpha v^\beta \partial_\beta w.$$

Similarly, $d_v(d_u w) = v^\alpha u^\beta \partial_\alpha \partial_\beta w + v^\alpha \partial_\alpha u^\beta \partial_\beta w$. The crucial point is that the second partial derivatives cancel each other, by (0.65).⁴⁰ Finally, we get

$$d_u(d_v w) - d_v(d_u w) = (u^\alpha \partial_\alpha v^\beta - v^\alpha \partial_\alpha u^\beta) \partial_\beta w = d_{[u, v]} w.$$

⁴⁰ It turns out that this is the main trick of Lie's approach to analysis and differential geometry. In addition, the tensorial property of the Riemann curvature tensor is based on a similar cancellation of second order partial derivatives.

This is (IV). Furthermore, the Leibniz rule (VII) follows from

$$d_v(\Theta w) = v^\alpha \partial_\alpha (\Theta w) = (v^\alpha \partial_\alpha \Theta) w + \Theta \cdot v^\alpha \partial_\alpha w = (d_v \Theta) w + \Theta d_v w.$$

□

In the present case, the Jacobi identity and the Bianchi identity are trivial consequences of the classical commutativity property (0.65) for partial derivatives. For more general situations, note the following. If $A, B, C : X \rightarrow X$ are linear operators on the linear space X , then cyclic permutation yields

$$\begin{aligned} & [A, [B, C]_-]_- + [B, [C, A]_-]_- + [C, [A, B]_-]_- \\ &= A(BC - CB) - (BC - CB)A + B(CA - AC) - (CA - AC)B \\ &+ C(AB - BA) - (AB - BA)C = 0. \end{aligned} \quad (0.66)$$

This means that the Jacobi identity is always satisfied for linear operators. Moreover, antisymmetrization yields

$$\begin{aligned} \text{Alt}(ABC) &:= \frac{1}{6}(ABC - ACB + BCA - BAC + CAB - CBA) \\ &= \frac{1}{6}(A[B, C]_- + B[C, A]_- + C[A, B]_-). \end{aligned}$$

The trouble with classical partial derivatives in the theory of general relativity. As a rule, the relations above are not invariant under general nonlinear coordinate transformations (i.e., by using local diffeomorphisms for local coordinates). In terms of physics, this means that the relations above do not possess any physical meaning.

As we will show, this lack of invariance can be cured by replacing partial derivatives by covariant partial derivatives.

Then the trivial Lie product does not vanish anymore, but the corresponding generalization determines the Riemann curvature tensor. The following sketched material will thoroughly be studied in Chap. 8.

The trouble with acceleration. Let $C : P = P(\sigma), \sigma_0 \leq \sigma \leq \sigma_1$, be a smooth curve on the space-time manifold \mathcal{M}^4 given by the equation $x = x(\sigma)$ with respect to a local coordinate system. As a rule, the equation

$$\ddot{x}^\gamma(\sigma) = 0, \quad \sigma_0 \leq \sigma \leq \sigma_1, \quad \gamma = 0, 1, 2, 3 \quad (0.67)$$

is not invariant under a change of local space-time coordinates. Set

$$\frac{D\dot{x}^\gamma(\sigma)}{d\sigma} := \ddot{x}^\gamma(\sigma) + \dot{x}^\alpha(\sigma) \Gamma_{\alpha\beta}^\gamma(x(\sigma)) \dot{x}^\beta(\sigma).$$

This is called the covariant derivative of $\sigma \mapsto \dot{x}(\sigma)$ with respect to the real parameter σ . It turns out that $\frac{D\dot{x}^\gamma(\sigma)}{d\sigma}$ is a tensorial family, in contrast to \ddot{x}^γ . The equation

$$\frac{D\dot{x}^\gamma(\sigma)}{d\sigma} = 0$$

describes geodesic lines. By the principle of index killing, the following vector functions possess an invariant meaning:

- $\mathbf{v}(\sigma) := \dot{x}^\gamma(\sigma) \partial_\gamma$ (velocity vector of the curve C),
- $\mathbf{a}(\sigma) := \frac{D\dot{x}^\gamma(\sigma)}{d\sigma} \partial_\gamma$ (acceleration vector of the curve C).

The partial covariant derivative as a key tool. We set

- $\nabla_\alpha v^\beta := \partial_\alpha v^\beta + \Gamma_{\alpha\lambda}^\beta v^\lambda$, and
- $\nabla_\alpha v_\beta = \partial_\alpha v_\beta - \Gamma_{\alpha\beta}^\lambda v_\lambda$.

More generally, we define

$$\nabla_\alpha T_{\beta_1 \dots \beta_n}^{\alpha_1 \dots \alpha_m} := \partial_\alpha T_{\beta_1 \dots \beta_n}^{\alpha_1 \dots \alpha_m} + \sum_{r=1}^m \Gamma_{\alpha\lambda}^{\alpha_r} T_{\beta_1 \dots \beta_n}^{\alpha_1 \dots \lambda \dots \alpha_m} - \sum_{s=1}^n \Gamma_{\alpha\beta_s}^\mu T_{\beta_1 \dots \mu \dots \beta_n}^{\alpha_1 \dots \alpha_m}.$$

Here, we replace the index α_r (resp. β_s) by λ (resp. μ). In addition, for an invariant function Θ , we define $\nabla_\alpha \Theta := \partial_\alpha \Theta$.

The covariant partial derivative ∇_α has the crucial property that it preserves tensorial families:

- If v^β is a tensorial family of type $(1, 0)$, then $\nabla_\alpha v^\beta$ is a tensorial family of type $(1, 1)$.
- If w_β is a tensorial family of type $(0, 1)$, then $\nabla_\alpha w_\beta$ is a tensorial family of type $(0, 2)$.
- If $T_{\beta_1 \dots \beta_n}^{\alpha_1 \dots \alpha_m}$ is a tensorial family of type (m, n) , then $\nabla_\alpha T_{\beta_1 \dots \beta_n}^{\alpha_1 \dots \alpha_m}$ is a tensorial family of type $(m, n + 1)$.

Let us mention two typical examples. Choose $\alpha, \beta, \gamma, \delta, \mu = 0, 1, 2, 3$.⁴¹ It follows from the key relation

$$\boxed{\nabla_\alpha (\nabla_\beta v^\delta) - \nabla_\beta (\nabla_\alpha v^\delta) = R_{\alpha\beta\gamma}^\delta v^\gamma} \quad (0.68)$$

that the Riemann curvature tensor measures the noncommutativity of the covariant partial derivative. The relation

$$\nabla_\mu R_{\alpha\beta\gamma}^\delta + \nabla_\alpha R_{\beta\mu\gamma}^\delta + \nabla_\beta R_{\mu\alpha\gamma}^\delta = 0 \quad (0.69)$$

represents the Bianchi identity in covariant formulation. This is based on cyclic permutation of the indices μ, α, β . The relation (0.69) is equivalent to

$$\boxed{\nabla_{[\mu} R_{\alpha\beta]\gamma}^\delta = 0} \quad (0.70)$$

which represents an antisymmetrization with respect to the indices μ, α, β . The Ricci identity reads as

$$R_{\alpha\beta\gamma}^\delta + R_{\beta\gamma\alpha}^\delta + R_{\gamma\alpha\beta}^\delta = 0.$$

This is equivalent to

$$\boxed{R_{[\alpha\beta\gamma]}^\delta = 0} \quad (0.71)$$

which represents an antisymmetrization with respect to the lower indices α, β, γ . Moreover, we have

$$\boxed{\nabla_\alpha g_{\beta\gamma} = 0.} \quad (0.72)$$

This is called the Ricci lemma. Consider the Bianchi identity (0.69). We will show later on that Ricci's lemma allows us to lower the index δ . This yields

⁴¹ By the Einstein convention, we sum over equal upper and lower indices from 0 to 3.

$$\nabla_\mu R_{\alpha\beta\gamma\delta} + \nabla_\alpha R_{\beta\mu\gamma\delta} + \nabla_\beta R_{\mu\alpha\gamma\delta} = 0 \quad (0.73)$$

based on cyclic permutation of the indices μ, α, β . Hence

$$\nabla_{[\mu} R_{\alpha\beta]\gamma\delta} = 0.$$

The very useful Ricci calculus principle of replacing partial derivatives by covariant partial derivatives. Consider the partial differential equation

$$\square\Theta = g^{\alpha\beta}\partial_\alpha\partial_\beta\Theta$$

in a fixed local coordinate system. For example, if $g_{\alpha\beta} = \eta_{\alpha\beta}$ (for all indices), and if we write ct, x, y, z instead of x^0, x^1, x^2, x^3 , respectively, then we obtain the wave equation

$$\square\Theta = \frac{1}{c^2} \frac{\partial^2\Theta}{\partial t^2} - \frac{\partial^2\Theta}{\partial x^2} - \frac{\partial^2\Theta}{\partial y^2} - \frac{\partial^2\Theta}{\partial z^2}. \quad (0.74)$$

It is our goal to rewrite (0.74) in such a way that it is valid in arbitrary local coordinate systems. To this end, we need an invariant expression which coincides with (0.74) in the special (x^0, x^1, x^2, x^3) -coordinate system chosen above with $g_{\alpha\beta} = \eta_{\alpha\beta}$. Since the Christoffel symbols vanish identically in this special coordinate system, we can replace the partial derivatives by the corresponding covariant partial derivatives. This yields the invariant formulation

$$\square\Theta = g^{\alpha\beta}\nabla_\alpha\nabla_\beta\Theta.$$

By the index killing principle, this expression is valid in each local coordinate system provided we use the tensorial family $g^{\alpha\beta}$. This simple trick can be used in order to write all the equations appearing in theoretical physics in such a way that they are valid in each coordinate system. For example, in Sect. 19.3.1 we will use this trick in order to formulate the Maxwell equations in electrodynamics with respect to an arbitrary space-time coordinate system. Observe that, for the Dirac equation of the relativistic electron, one has to replace tensorial families by spinorial families (see Vol. IV).

The Lie Structure behind Curvature

Let the symbol $\text{Vect}(\mathcal{M}^4)$ denote the space of all smooth velocity vector fields \mathbf{v} on the space-time manifold \mathcal{M}^4 . With respect to local coordinates, we write $\mathbf{v} = v^\alpha\partial_\alpha$. In what follows, let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \text{Vect}(\mathcal{M}^4)$. Sophus Lie frequently used the fact that

$$u^\alpha\partial_\alpha v^\beta - v^\alpha\partial_\alpha u^\beta$$

forms a tensorial family on \mathcal{M}^4 . By index killing, we get the invariant expression

$$\mathcal{L}_{\mathbf{u}}\mathbf{v} = (u^\alpha\partial_\alpha v^\beta - v^\alpha\partial_\alpha u^\beta)\partial_\beta.$$

This is called the Lie derivative of the velocity field \mathbf{v} with respect to the velocity field \mathbf{u} .

The Lie algebra $\text{Vect}(\mathcal{M}^4)$. Using the Lie product

$$[\mathbf{u}, \mathbf{v}] := \mathcal{L}_{\mathbf{u}}\mathbf{v},$$

the linear space $\text{Vect}(\mathcal{M}^4)$ of smooth velocity vector fields on \mathcal{M}^4 becomes a real Lie algebra. Explicitly, this means that for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \text{Vect}(\mathcal{M}^4)$ and all real numbers λ, μ , we have:

- $[\lambda \mathbf{u} + \mu \mathbf{v}, \mathbf{w}] = \lambda [\mathbf{u}, \mathbf{w}] + \mu [\mathbf{v}, \mathbf{w}]$ (distributivity),
- $[\mathbf{v}, \mathbf{w}] = -[\mathbf{w}, \mathbf{v}]$ (antisymmetry),
- $[\mathbf{u}, [\mathbf{v}, \mathbf{w}]] + [\mathbf{v}, [\mathbf{w}, \mathbf{u}]] + [\mathbf{w}, [\mathbf{u}, \mathbf{v}]] = 0$ (Jacobi identity).

The physical meaning of the Lie derivative in terms of the flow of fluid particles will be considered in Sect. 11.2.

The covariant differential of a velocity field. For $\mathbf{u}, \mathbf{v} \in \text{Vect}(\mathcal{M}^4)$, we have the tensorial family $\nabla_\alpha v^\beta$. Using the index killing principle, we get the invariant expression

$$D\mathbf{v} := (\nabla_\alpha v^\beta) dx^\alpha \otimes \partial_\beta.$$

This tensor field of type (1, 1) on the space-time manifold \mathcal{M}^4 is called the covariant differential of the velocity vector field \mathbf{v} . Furthermore, we set

$$D_{\mathbf{u}}\mathbf{v} := (D\mathbf{v})(\mathbf{u}).$$

This is called the covariant directional derivative of the velocity vector field \mathbf{v} on \mathcal{M}^4 with respect to the velocity vector field \mathbf{u} . Explicitly,

$$D_{\mathbf{u}}\mathbf{v} = (u^\alpha \nabla_\alpha v^\beta) \partial_\beta.$$

Two key relations for velocity vector fields. For all smooth velocity vector fields $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \text{Vect}(\mathcal{M}^4)$ on the space-time manifold \mathcal{M}^4 , we have

$$[\mathbf{u}, \mathbf{v}] = D_{\mathbf{u}}\mathbf{v} - D_{\mathbf{v}}\mathbf{u}, \quad (0.75)$$

and

$$\mathbf{F}(\mathbf{u}, \mathbf{v})\mathbf{w} = (D_{\mathbf{u}}D_{\mathbf{v}} - D_{\mathbf{v}}D_{\mathbf{u}} - D_{[\mathbf{u}, \mathbf{v}]})\mathbf{w}. \quad (0.76)$$

In terms of geometry, the Riemann curvature operator $\mathbf{F}(\mathbf{u}, \mathbf{v})$ was introduced on page 33 by using parallel transport of a velocity vector along a sufficiently small closed path. In terms of analysis, relation (0.76) connects the Riemann curvature operator with the covariant directional derivative. Equivalently, relation (0.75) reads as

$$\mathcal{L}_{\mathbf{u}}\mathbf{v} = D_{\mathbf{u}}\mathbf{v} - D_{\mathbf{v}}\mathbf{u}.$$

This connects the Lie derivative with the covariant directional derivative. In terms of the covariant directional derivative $D_{\mathbf{u}}\mathbf{v}$, we get the following two key formulas for fixed smooth velocity vector fields \mathbf{u}, \mathbf{v} on the 4-dimensional space-time manifold \mathcal{M}^4 equipped with the metric tensor \mathbf{g} :

$$D_{\mathbf{u}}\mathbf{v} - D_{\mathbf{v}}\mathbf{u} - [\mathbf{u}, \mathbf{v}] = 0 \quad (0.77)$$

and

$$\mathbf{F}(\mathbf{u}, \mathbf{v}) = D_{\mathbf{u}}D_{\mathbf{v}} - D_{\mathbf{v}}D_{\mathbf{u}} - D_{[\mathbf{u}, \mathbf{v}]}, \quad (0.78)$$

where $\mathbf{F}(\mathbf{u}, \mathbf{v})$ represents the Riemann curvature operator.⁴² Introducing Weyl's torsion operator

$$\mathbf{T}(\mathbf{u}, \mathbf{v}) := D_{\mathbf{u}}\mathbf{v} - D_{\mathbf{v}}\mathbf{u} - [\mathbf{u}, \mathbf{v}],$$

⁴² Gauss (1777–1855), Riemann (1826–1866), Lie (1842–1899), Ricci-Curbastro (1853–1925), Élie Cartan (1869–1951), Levi-Civita (1873–1941), Einstein (1879–1955), Weyl (1885–1955).

the first key relation (0.77) reads as

$$\mathbf{T}(\mathbf{u}, \mathbf{v}) = 0.$$

That is, the torsion vanishes in Riemannian and pseudo-Riemannian geometry. But note that there are more general geometries where the torsion does not vanish. It turns out that the two key relations (0.77) and (0.78) govern Riemannian and pseudo-Riemannian geometry. If the metric tensor \mathbf{g} is constant on \mathcal{M}^4 , then we have the situation of Einstein's theory of special relativity where the gravitational force vanishes.⁴³ Then the two key relations (0.77) and (0.78) pass over to the two classical relations

- $d_{\mathbf{u}}\mathbf{v} - d_{\mathbf{v}}\mathbf{u} - [\mathbf{u}, \mathbf{v}] = 0$, and
- $d_{\mathbf{u}}d_{\mathbf{v}} - d_{\mathbf{v}}d_{\mathbf{u}} - d_{[\mathbf{u}, \mathbf{v}]} = 0$,

respectively. Here, $d_{\mathbf{u}}\mathbf{v}$ denotes the classical directional derivative. Relation (0.76) allows far-reaching generalizations which represent the fundamental principle

$$\text{force} = \text{curvature}$$

in modern physics. This will be studied in the present volume.

Parallel Transport and the Covariant Directional Derivative

We are given the smooth curve

$$C : P = P(\sigma), \quad \sigma_0 < \sigma < \sigma_0. \quad (0.79)$$

Let $\sigma_0 < 0 < \sigma_1$. Set $P_0 := P(0)$. We want to characterize both the covariant directional derivative $D_{\mathbf{u}}\mathbf{v}(P_0)$ and the covariant derivative $\frac{D\mathbf{v}(0)}{d\sigma}$ with respect to the real parameter σ by a limiting process based on parallel transport.

Covariant directional derivative. For a smooth classical real-valued function $f : \mathbb{R} \rightarrow \mathbb{R}$, the derivative $\dot{f}(0)$ at the point $\sigma = 0$ is given by

$$\dot{f}(0) = \lim_{\sigma \rightarrow 0} \frac{f(\sigma) - f(0)}{\sigma}.$$

As we will show later on, the generalization to the covariant directional derivative reads as follows:

$$\boxed{D_{\mathbf{u}}\mathbf{v}(P_0) = \lim_{\sigma \rightarrow 0} \frac{\Pi_{\sigma}^{-1}\mathbf{v}(P(\sigma)) - \mathbf{v}(P_0)}{\sigma}.} \quad (0.80)$$

Let us discuss this. We are given the smooth velocity vector field $\mathbf{v} = \mathbf{v}(P)$ on the space-time manifold \mathcal{M}^4 . Fix the point $P_0 \in \mathcal{M}^4$, and fix the tangent vector $\mathbf{u} \in T_{P_0}\mathcal{M}^4$ at the point P_0 . We choose a smooth curve C as given by (0.79) which passes through the point P_0 at the parameter value $\sigma = 0$ and has the tangent vector \mathbf{u} at P_0 . With respect to a local coordinate system, the curve C is given by $x = x(\sigma)$ where $x(0)$ corresponds to the point P_0 and $\mathbf{u} = \dot{x}^\alpha(0)\partial_\alpha$. Let us now consider the parallel transport along the curve C . To this end, we introduce the operator

$$\Pi_\sigma : T_{P_0}\mathcal{M}^4 \rightarrow T_{P(\sigma)}\mathcal{M}^4.$$

⁴³ In this special case, the Christoffel symbols vanish, and the covariant partial derivative coincides with the classical partial derivative, $\nabla_\alpha = \partial_\alpha$.

By definition, the tangent vector $\Pi_\sigma \mathbf{w}$ (at the point $P(\sigma)$) is obtained from the tangent vector \mathbf{w} (at the point P_0) by using parallel transport from the point P_0 to the point $P(\sigma)$ along the curve C . This parallel transport is reversed by the inverse operator

$$\Pi_\sigma^{-1} : T_{P(\sigma)}\mathcal{M}^4 \rightarrow T_{P_0}\mathcal{M}^4.$$

This completes the explanation of the notation used in (0.80).

The geometric intuition behind (0.80) will be explained in Chap. 9.5 by considering the situation on a sphere. Naively, one would use the limit

$$\lim_{\sigma \rightarrow 0} \frac{\mathbf{v}(P(\sigma)) - \mathbf{v}(P_0)}{\sigma}.$$

This expression can be computed by using local coordinates. However, it turns out that, as a rule, the result depends on the choice of the local coordinates, that is, the expression does not possess any invariant geometric (or physical) meaning. Roughly speaking, the reason for this is the fact that the tangent vectors $\mathbf{v}(P_0)$ and $\mathbf{v}(P(\sigma))$ live in different tangent spaces. In order to be able to compute the vector difference in the same tangent space $T_{P_0}\mathcal{M}^4$, we replace the tangent vector $\mathbf{v}(P(\sigma))$ at the point $P(\sigma)$ by the parallel transported tangent vector $\Pi_\sigma^{-1}\mathbf{v}(P(\sigma))$ at the point P_0 .

From the physical point of view, it is impossible to compare physical quantities (e.g., fields) at different space-time points without using the transport of physical information.

This observation is crucial for gauge theory in modern physics (the theory of general relativity and the Standard Model in particle physics).

Covariant derivatives with respect to the real parameter σ . Let the family $\mathbf{v} = \mathbf{v}(\sigma)$ of velocity vectors be given along the curve C from (0.79), that is, the vector $\mathbf{v}(\sigma)$ lives in the tangent space $T_{P(\sigma)}\mathcal{M}^4$ for all $\sigma \in]\sigma_0, \sigma_1[$. We define

$$\boxed{\frac{D\mathbf{v}(0)}{d\sigma} := \lim_{\sigma \rightarrow 0} \frac{\Pi_\sigma^{-1}(\mathbf{v}(\sigma)) - \mathbf{v}(P_0)}{\sigma}}. \quad (0.81)$$

Similarly, we define the covariant derivative $\frac{D\mathbf{v}(\sigma)}{d\sigma}$ at the parameter $\sigma \in]\sigma_0, \sigma_1[$. With respect to local coordinates, we have

$$\mathbf{v}(\sigma) = v^\gamma(\sigma)\partial_\gamma, \quad \frac{D\mathbf{v}(\sigma)}{d\sigma} = \frac{Dv^\gamma(\sigma)}{d\sigma} \partial_\gamma, \quad \sigma \in]\sigma_0, \sigma_1[$$

together with

$$\frac{Dv^\gamma(\sigma)}{d\sigma} := \dot{v}^\gamma(\sigma) + \dot{x}^\alpha(\sigma)\Gamma_{\alpha\beta}^\gamma(x(\sigma)) \cdot v^\beta(\sigma), \quad \gamma = 0, 1, 2, 3.$$

Here, the curve C corresponds to $x^\gamma = x^\gamma(\sigma)$ with $\sigma_0 < \sigma < \sigma_1$. By definition, the family $\mathbf{v} = \mathbf{v}(\sigma)$ of velocity vectors is parallel along the curve C iff

$$\boxed{\frac{D\mathbf{v}(\sigma)}{d\sigma} = 0 \quad \text{for all } \sigma \in]\sigma_0, \sigma_1[.}$$

As we will show later on, this ordinary differential equation describes the transport of physical information in gauge theory (theory of general relativity and the Standard Model in physics).

Parallel transport respects the inner product. Suppose that the two smooth velocity vector fields $\mathbf{v} = \mathbf{v}(\sigma)$ and $\mathbf{w} = \mathbf{w}(\sigma)$ are given along the curve

$C : P = P(\sigma)$, $\sigma_0 < \sigma < \sigma_1$. Then, for all $\sigma \in]\sigma_0, \sigma_1[$, we get the following generalized Leibniz product rule:

$$\frac{d}{d\sigma} \langle \mathbf{v}(\sigma) | \mathbf{w}(\sigma) \rangle = \left\langle \frac{D\mathbf{v}(\sigma)}{d\sigma} | \mathbf{w}(\sigma) \right\rangle + \left\langle \mathbf{v}(\sigma) | \frac{D\mathbf{w}(\sigma)}{d\sigma} \right\rangle. \quad (0.82)$$

This implies that, for all smooth velocity vector fields $\mathbf{u}, \mathbf{v}, \mathbf{w}$ on the space-time manifold \mathcal{M}^4 , we have

$$d_{\mathbf{u}} \langle \mathbf{v} | \mathbf{w} \rangle = \langle D_{\mathbf{u}} \mathbf{v} | \mathbf{w} \rangle + \langle \mathbf{v} | D_{\mathbf{u}} \mathbf{w} \rangle.$$

Mnemonicly, we write

$$d \langle \mathbf{v} | \mathbf{w} \rangle = \langle D\mathbf{v} | \mathbf{w} \rangle + \langle \mathbf{v} | D\mathbf{w} \rangle.$$

In particular, if $\mathbf{v} = \mathbf{v}(\sigma)$ and $\mathbf{w} = \mathbf{w}(\sigma)$ are parallel along the curve C , then $\frac{D\mathbf{v}(\sigma)}{d\sigma} = \frac{D\mathbf{w}(\sigma)}{d\sigma} \equiv 0$. By (0.82),

$$\langle \mathbf{v}(\sigma) | \mathbf{w}(\sigma) \rangle = \text{const} \quad \text{for all } \sigma \in]\sigma_0, \sigma_1[.$$

This means that the parallel transport of tangent vectors (i.e., velocity vectors) on the 4-dimensional space-time manifold preserves the (indefinite) inner product.

The covariant differential for general tensor fields. Later on, starting from the covariant differential $D\mathbf{v}$ of the velocity vector field \mathbf{v} , we will introduce the covariant differential $D\mathbf{T}$ for general smooth tensor fields \mathbf{T} by using

- $(D\omega)(\mathbf{v}) = \omega(D\mathbf{v})$ (duality between vector fields \mathbf{v} and covector fields ω), and
- $D(\mathbf{T} \otimes \mathbf{S}) = D\mathbf{T} \otimes \mathbf{S} + \mathbf{T} \otimes D\mathbf{S}$ (the Leibniz product rule for tensor fields \mathbf{T} and \mathbf{S}).

Explicitly, for the given tensor field

$$\mathbf{T} = T_{\alpha_1 \dots \alpha_n}^{\beta_1 \dots \beta_m} dx^{\alpha_1} \otimes \dots \otimes dx^{\alpha_n} \otimes \partial_{\beta_1} \otimes \dots \otimes \partial_{\beta_m},$$

we obtain

$$D\mathbf{T} = \nabla_{\alpha} T_{\alpha_1 \dots \alpha_n}^{\beta_1 \dots \beta_m} dx^{\alpha} \otimes dx^{\alpha_1} \otimes \dots \otimes dx^{\alpha_n} \otimes \partial_{\beta_1} \otimes \dots \otimes \partial_{\beta_m}.$$

The Generalized Riemann Curvature Tensor in Modern Mathematics and Physics

Let $GL(N, \mathbb{C})$ denote the Lie group of all invertible complex $(N \times N)$ -matrices. This is a real manifold of dimension $2N^2$. Moreover, let $gl(N, \mathbb{C})$ denote the Lie algebra to the Lie group $GL(N, \mathbb{C})$. Explicitly, the real Lie algebra $gl(N, \mathbb{C})$ consists of all complex $(N \times N)$ -matrices. The following generalization of gauge theory is crucial for the Standard Model in particle physics. Note the following:

The structural equation

$$\mathcal{F}_{\alpha\beta} = \partial_{\alpha}\mathcal{A}_{\beta} - \partial_{\beta}\mathcal{A}_{\alpha} + \mathcal{A}_{\alpha}\mathcal{A}_{\beta} - \mathcal{A}_{\beta}\mathcal{A}_{\alpha} \quad (0.83)$$

with $\alpha, \beta = 0, 1, 2, 3$ makes sense if $\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ are complex $(N \times N)$ -matrices with $N = 1, 2, \dots$

This means that the matrices $\mathcal{A}_0, \mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3$ are contained in the real Lie algebra $gl(N, \mathbb{C})$. More generally, let the symbol \mathcal{G} denote a closed subgroup of the Lie group $GL(N, \mathbb{C})$.

Then \mathcal{G} is a Lie subgroup of $GL(N, \mathbb{C})$.

Furthermore, let \mathcal{LG} denote the real Lie algebra to the Lie group \mathcal{G} . Choose

$$\mathcal{A}_\alpha \in \mathcal{LG}, \quad \alpha = 0, 1, 2, 3.$$

Then, the Lie product $[A_\alpha, A_\beta] := \mathcal{A}_\alpha \mathcal{A}_\beta - \mathcal{A}_\beta \mathcal{A}_\alpha$ is also contained in the Lie algebra \mathcal{LG} , and hence $\mathcal{F}_{\alpha\beta}$ is contained in the Lie algebra \mathcal{LG} . That is, both

- the connection form $\mathcal{A} = \mathcal{A}_\alpha dx^\alpha$ and
- the curvature form $\mathcal{F} = \frac{1}{2} \mathcal{F}_{\alpha\beta} dx^\alpha \wedge dx^\beta$

are differential forms with values in the Lie algebra \mathcal{LG} . Here, \mathcal{G} is called the gauge group, and \mathcal{LG} is called the gauge Lie algebra. For example, choose

$$\mathcal{G} = SU(N) \quad \text{and} \quad \mathcal{LG} = su(N), \quad N = 2, 3, \dots$$

Here, the Lie group $SU(N)$ consists of all complex $(N \times N)$ -matrices A with $AA^\dagger = I$ and the determinant $\det(A) = 1$. This is called the special unitary group (on the complex Hilbert space \mathbb{C}^N). The real Lie algebra $su(N)$ to the Lie group $SU(N)$ consists of all complex $(N \times N)$ -matrices with $A + A^\dagger = 0$ and the trace $\text{tr}(A) = 0$. Passing to components, we get

- $\mathcal{A}_\alpha = (\gamma_{\alpha L}^K)$, and
- $\mathcal{F}_{\alpha\beta} = (r_{L\alpha\beta}^K)$.

For the indices, we have $\alpha, \beta = 0, 1, 2, 3$, and $K, L = 1, \dots, N$. Here, the functions $\gamma_{\alpha L}^K$ (resp. $r_{L\alpha\beta}^K$) are called the generalized Christoffel symbols (resp. the components of the generalized Riemann curvature tensor).

Cartan's global approach. As we will show later on, it turns out that, as a rule, the connection form \mathcal{A} and the curvature form \mathcal{F} are not invariant under local gauge transformations. Therefore, Cartan introduced

- the global connection form A with values in the Lie algebra \mathcal{LG} , and
- the global curvature form F with values in the Lie algebra \mathcal{LG}

on an appropriate principal fiber bundle $\mathcal{P}(\mathcal{M}^4)$ over the space-time manifold \mathcal{M}^4 . Then Cartan's local structural equation (0.83) is generalized to the global structural equation

$$\boxed{F = DA.} \tag{0.84}$$

In terms of physics, this global structural equation represents the most elegant mathematical formulation of the principle “force equals curvature.” In terms of mathematics, equation (0.84) represents a far-reaching generalization of Gauss' theorema egregium. Every differential form F represented by (0.84) satisfies the integrability condition

$$\boxed{DF = 0} \tag{0.85}$$

which is called the (global) Bianchi identity (see Chap. 17).

Parallel Transport of Physical Information and the Local Phase Factor

Let us sketch the basic ideas. We are given the curve $x = x(\sigma)$, $\sigma_0 \leq \sigma \leq \sigma_1$, on the space-time manifold \mathcal{M}^4 . The ordinary differential equation

$$\boxed{\dot{G}(\sigma) = -\dot{x}^\alpha(\sigma)\mathcal{A}_\alpha(x(\sigma)) \cdot G(\sigma)} \quad (0.86)$$

with the initial condition $G(\sigma_0) = I$ (unit matrix) is called the differential equation of parallel transport (for the phase factor G). Since $\mathcal{A}_\alpha(P)$ is contained in the Lie algebra \mathcal{LG} for all indices α and all points P of the space-time manifold, the solution $\sigma \mapsto G(\sigma)$ has the property that $G(\sigma)$ is an element of the Lie group \mathcal{G} for all values of the parameter σ . As we will show later on, the differential equation (0.86) can be used in order to define the covariant differential DA by means of the classical differential dA and a suitable projection defined on the tangent spaces of the appropriate principal fiber bundle (horizontal tangent vectors; see Chap. 17).

The local phase factor of a physical field ψ . Let

$$\psi(x) = \begin{pmatrix} \psi^1(x) \\ \vdots \\ \psi^N(x) \end{pmatrix}$$

be a complex column matrix with N rows which depends on the space-time point x . In terms of physics, the function $x \mapsto \psi(x)$ describes a physical field (e.g., the wave function of an electron in the Standard Model in particle physics). Let $\sigma \mapsto G(\sigma)$ be the unique solution of the differential equation (0.86) of parallel transport. By definition,

$$\psi(\sigma) := G(\sigma)\psi(x(\sigma_0)), \quad \sigma_1 \leq \sigma \leq \sigma_1. \quad (0.87)$$

This equation describes the parallel transport of the physical field ψ along the curve $x = x(\sigma)$, $\sigma_0 \leq \sigma \leq \sigma_1$. Here, $G(\sigma)$ is called the phase factor of the physical field ψ at the space-time point $x(\sigma)$. Using this terminology, the equation (0.86) describes the parallel transport of the local phase function $\sigma \mapsto G(\sigma)$. Differentiating equation (0.87) with respect to the parameter σ , we get the following differential equation of parallel transport for the physical field ψ :

$$\boxed{\dot{\psi}(\sigma) = -\dot{x}^\alpha(\sigma)\mathcal{A}_\alpha(x(\sigma)) \psi(\sigma), \quad \sigma_0 \leq \sigma \leq \sigma_1.}$$

In terms of physics, this differential equation describes the transport of physical information via phase factor.

The Modern Language of Fiber Bundles in Mathematics and Physics

Physical fields are sections of fiber bundles. The qualitative (i.e., topological) structure of physical fields is determined by the topological structure of the corresponding fiber bundles. The prototype of a fiber bundle is the tangent bundle of a manifold.

Folklore

The tangent bundle and tangent vector fields. Consider a tangent vector field $\mathbf{v} = \mathbf{v}(P)$ on the space-time manifold \mathcal{M}^4 . Such a field assigns to each point P of the manifold \mathcal{M}^4 the tangent vector $\mathbf{v}(P)$ at the point P , that is, $\mathbf{v}(P) \in T_P\mathcal{M}^4$. In order to handle conveniently tangent vector fields as mathematical objects, we will describe them as maps of the form

$$\boxed{s : \mathcal{M}^4 \rightarrow T\mathcal{M}^4}, \quad (0.88)$$

which are called sections s of the tangent bundle $T\mathcal{M}^4$. Let us discuss this basic notion in modern mathematics. By definition, the tangent bundle $T\mathcal{M}^4$ of the manifold \mathcal{M}^4 consists of all the ordered pairs

$$(P, \mathbf{v})$$

where P is an arbitrary point of \mathcal{M}^4 , and \mathbf{v} is an arbitrary tangent vector of \mathcal{M}^4 at the point P . Briefly,

$$T\mathcal{M}^4 := \{(P, \mathbf{v}) : P \in \mathcal{M}^4, \mathbf{v} \in T_P\mathcal{M}^4\}.$$

The map $s : \mathcal{M}^4 \rightarrow T\mathcal{M}^4$ is called a section of the tangent bundle $T\mathcal{M}^4$ iff the image $s(P)$ has the form $(P, \mathbf{v}(P))$ with $\mathbf{v}(P) \in T_P\mathcal{M}^4$ for all points $P \in \mathcal{M}^4$.

Let us add some more terminology used in modern mathematics. Setting $\pi(P, \mathbf{v}) := P$, we get the so-called projection map

$$\pi : T\mathcal{M}^4 \rightarrow \mathcal{M}^4.$$

The pre-image $\mathcal{F}_P := \pi^{-1}(P)$ is called the fiber of the tangent bundle $T\mathcal{M}^4$ over the base point P . Explicitly,

$$\mathcal{F}_P := \{(P, \mathbf{v}) : \mathbf{v} \in T_P\mathcal{M}^4\}.$$

There exists a one-to-one correspondence $T_P\mathcal{M}^4 \leftrightarrow \mathcal{F}_P$. Therefore, the tangent spaces of the manifold \mathcal{M}^4 can be identified with the fibers of the tangent bundle $T\mathcal{M}^4$. If $P \neq Q$, then $\mathcal{F}_P \cap \mathcal{F}_Q = \emptyset$. We have

$$T\mathcal{M}^4 = \bigcup_{P \in \mathcal{M}^4} \mathcal{F}_P.$$

That is, the tangent bundle $T\mathcal{M}^4$ is the union of the pairwise disjoint fibers \mathcal{F}_P . Observe that the map $s : \mathcal{M}^4 \rightarrow T\mathcal{M}^4$ is a section of the tangent bundle iff the following diagram is commutative:

$$\begin{array}{ccc} \mathcal{M}^4 & \xrightarrow{s} & T\mathcal{M}^4 \\ & \searrow \text{id} & \downarrow \pi \\ & & \mathcal{M}^4. \end{array} \quad (0.89)$$

That is, $\pi(s(P)) = P$ for all $P \in \mathcal{M}^4$. In other words, the section s respects fibers, that is, $s(P)$ lives in the fiber \mathcal{F}_P for all $P \in \mathcal{M}^4$. Synonymously, we will use the notions

- tangent vector and
- velocity vector.

In particular, tangent vector fields are also called velocity vector fields. In a quite natural way, we are able to introduce local coordinates on the tangent bundle. To this end, we choose an arbitrary local coordinate system of the basis manifold \mathcal{M}^4 . We assign to the point P of \mathcal{M}^4 the local coordinate $x = (x^0, x^1, x^2, x^3)$. Moreover, we have $\mathbf{v} = v^\alpha \partial_\alpha$. Finally, we assign to the point (P, \mathbf{v}) of the tangent bundle $T\mathcal{M}^4$ the local coordinates

$$(x^0, x^1, x^2, x^3, v^0, v^1, v^2, v^3).$$

This way, the tangent bundle $T\mathcal{M}^4$ becomes a real 8-dimensional manifold. By definition, the velocity vector field $\mathbf{v} = \mathbf{v}(P)$ is smooth on \mathcal{M}^4 iff the corresponding section $s : \mathcal{M}^4 \rightarrow T\mathcal{M}^4$ is a smooth map between the two manifolds \mathcal{M}^4 and $T\mathcal{M}^4$. The symbol $\text{Vect}(\mathcal{M}^4)$ denotes the set of all smooth tangent vector fields on \mathcal{M}^4 .

The cotangent bundle and cotangent vector fields. Now let us pass from tangent vectors to the dual objects called cotangent vectors. Precisely the linear functionals

$$\omega : T_P\mathcal{M}^4 \rightarrow \mathbb{R}$$

on the tangent space $T_P\mathcal{M}^4$ are called cotangent vectors of the manifold \mathcal{M}^4 at the point P . By definition, the set of all these linear functionals forms the dual space $T_P^*\mathcal{M}^4$ to the tangent space $T_P\mathcal{M}^4$. This dual space $T_P^*\mathcal{M}^4$ is called the cotangent space of the manifold \mathcal{M}^4 at the point P . We have $\omega \in T_P^*\mathcal{M}^4$ iff there exist real numbers $\omega_0, \omega_1, \omega_2, \omega_3$ such that

$$\omega = \omega_\alpha dx^\alpha.$$

Recall that $dx^\alpha(\mathbf{v}) := v^\alpha$ for all tangent vectors $\mathbf{v} = v^\alpha \partial_j$ at the point P . A cotangent vector field

$$\omega = \omega(P) \quad \text{on } \mathcal{M}^4$$

assigns to each point P of \mathcal{M}^4 the cotangent vector $\omega(P)$. By definition, the cotangent bundle $T^*\mathcal{M}^4$ consists of all ordered pairs

$$(P, \omega)$$

where P is an arbitrary point of \mathcal{M}^4 , and ω is an arbitrary cotangent vector of \mathcal{M}^4 at the point P . Briefly,

$$T^*\mathcal{M}^4 := \{(P, \omega) : P \in \mathcal{M}^4, \omega \in T_P^*(\mathcal{M}^4)\}.$$

The map

$$s : \mathcal{M}^4 \rightarrow T^*\mathcal{M}^4$$

is called a section of the cotangent bundle $T^*\mathcal{M}^4$ iff the image $s(P)$ has the form $(P, \omega(P))$ with $\omega(P) \in T_P^*\mathcal{M}^4$ for all points $P \in \mathcal{M}^4$. Setting $\pi(P, \omega) := P$, we get the so-called projection map

$$\pi : T^*\mathcal{M}^4 \rightarrow \mathcal{M}^4.$$

For the pre-image, we obtain $\pi^{-1}(P) = \{P\} \times T_P^*\mathcal{M}^4$. This is called the fiber \mathcal{F}_P over the base point P . Explicitly,

$$\mathcal{F}_P = \{(P, \omega) : P \in \mathcal{M}^4, \omega \in T_P^*\mathcal{M}^4\}.$$

Therefore, the cotangent spaces of the manifold \mathcal{M}^4 can be identified with the fibers of the cotangent bundle $T^*\mathcal{M}^4$. The map

$$s : \mathcal{M}^4 \rightarrow T^*\mathcal{M}^4$$

is called a section of the cotangent bundle iff the following diagram is commutative:

$$\begin{array}{ccc}
 \mathcal{M}^4 & \xrightarrow{s} & T^*\mathcal{M}^4 \\
 & \searrow \text{id} & \downarrow \pi \\
 & & \mathcal{M}^4.
 \end{array} \tag{0.90}$$

That is, $\pi(s(P)) = P$ for all $P \in \mathcal{M}^4$. Sections of the cotangent bundle $T^*\mathcal{M}^4$ correspond to cotangent vector fields $\omega = \omega(P)$ on \mathcal{M}^4 . Synonymously, we use the following notions:

- cotangent vector,
- velocity covector (or briefly covector),
- differential 1-form.

In particular, cotangent vector fields are also called velocity covector fields (or fields of differential 1-forms). Using $\omega = \omega_\alpha dx^\alpha$, we assign to the point (P, ω) of the cotangent bundle $T^*\mathcal{M}^4$ the local coordinates

$$(x^0, x^1, x^2, x^3, \omega_0, \omega_1, \omega_2, \omega_3).$$

This way, the cotangent bundle $T^*\mathcal{M}^4$ becomes a real 8-dimensional manifold. By definition, the cotangent vector field $\omega = \omega(P)$ is smooth on \mathcal{M}^4 iff the corresponding section $s : \mathcal{M}^4 \rightarrow T^*\mathcal{M}^4$ is a smooth map. The symbol $\text{Covect}(\mathcal{M}^4)$ denotes the set of all smooth cotangent vector fields on \mathcal{M}^4 .

Tensor bundles and tensor fields. By definition, a tensor of type $(0, 2)$ of the manifold \mathcal{M}^4 at the point P is a bilinear map of the form

$$\mathbf{g} : T_P^*\mathcal{M}^4 \times T_P^*\mathcal{M}^4 \rightarrow \mathbb{R}.$$

This means that $\mathbf{g} = g_{\alpha\beta} dx^\alpha \otimes dx^\beta$ where $g_{\alpha\beta}$ are fixed, but otherwise arbitrary real numbers for all indices $\alpha, \beta = 0, 1, 2, 3$. The tensor bundle $T_2^0(\mathcal{M}^4)$ consists of all the ordered pairs

$$(P, \mathbf{g})$$

where P is an arbitrary point of \mathcal{M}^4 , and \mathbf{g} is an arbitrary tensor of type $(0, 2)$. The map $s : \mathcal{M}^4 \rightarrow T_2^0\mathcal{M}^4$ is called a section of the tensor bundle $T_2^0\mathcal{M}^4$ iff the following diagram is commutative:

$$\begin{array}{ccc}
 \mathcal{M}^4 & \xrightarrow{s} & T_2^0\mathcal{M}^4 \\
 & \searrow \text{id} & \downarrow \pi \\
 & & \mathcal{M}^4.
 \end{array} \tag{0.91}$$

Here, $\pi(P, \mathbf{g}) := P$. Sections of the tensor bundle $T_2^0\mathcal{M}^4$ correspond to tensor fields $\mathbf{g} = \mathbf{g}(P)$ of type $(0, 2)$ on \mathcal{M}^4 . Using $\mathbf{g} = g_{\alpha\beta} dx^\alpha \otimes dx^\beta$, we assign to the point (P, \mathbf{g}) of the tensor bundle $T_2^0\mathcal{M}^4$ the local coordinates

$$(x^\gamma, g_{\alpha\beta})_{\alpha, \beta, \gamma=0,1,2,3}.$$

This way, the tensor bundle $T_2^0\mathcal{M}^4$ becomes a real 20-dimensional manifold. The symbol $\otimes_2^0(\mathcal{M}^4)$ denotes the set of all smooth sections $s : \mathcal{M}^4 \rightarrow T_2^0\mathcal{M}^4$ of the tensor bundle $T_2^0\mathcal{M}^4$ (i.e., the set of all smooth tensor fields of type $(0, 2)$ on the space-time manifold \mathcal{M}^4).

Analogously, we will introduce the tensor bundle $T_n^m \mathcal{M}^4$, and the set $\bigotimes_n^m(\mathcal{M}^4)$ of all smooth tensor fields of type (m, n) on \mathcal{M}^4 .

The symbol $\bigwedge^m(\mathcal{M}^4)$ denotes the set of all smooth antisymmetric tensor fields of type $(m, 0)$ on \mathcal{M}^4 . As we will show later on, such tensor fields coincide with fields of differential m -forms on \mathcal{M}^4 .

The topological structure of velocity vector fields and characteristic classes. It turns out that the qualitative behavior of physical fields depends on the topological properties of the corresponding fiber bundles. In this connection, characteristic classes of fiber bundles are the most important topological invariants which govern the topological structure of physical fields. To illustrate this, let us consider the case of velocity vector fields $\mathbf{v} = \mathbf{v}(P)$ on the n -dimensional unit sphere \mathbb{S}^n . We want to study the following three basic questions:

- the existence of stagnation points,
- the maximal number of linearly independent continuous velocity vector fields (in the global sense), and
- the existence of a global parallel transport.

By definition, the velocity vector field $\mathbf{v} = \mathbf{v}(P)$ has the critical (or stagnation) point P_0 iff $\mathbf{v}(P_0) = 0$.

Stagnation points. The Euler characteristic of the sphere \mathbb{S}^n is given by

$$\chi(\mathbb{S}^n) = 1 + (-1)^n, \quad n = 1, 2, \dots$$

Let n be even. There exists a smooth field $\omega = \omega(P)$ of differential n -forms on the sphere \mathbb{S}^n such that we have the integral representation

$$\chi(\mathbb{S}^n) = \int_{\mathbb{S}^n} \omega, \quad n = 2, 4, 6, \dots$$

Physicists call the Euler characteristic $\chi(\mathbb{S}^n)$ a topological charge. By definition, the Euler class $[\omega]$ of the tangent bundle $T\mathbb{S}^n$ is the set

$$[\omega] := \{\omega + d\nu\}$$

where ν is an arbitrary smooth differential $(n-1)$ -form on \mathbb{S}^n . In other words,

$$[\omega] \in H^n(\mathbb{S}^n), \quad n = 1, 2, \dots$$

That is, the Euler class $[\omega]$ is an element of $H^n(\mathbb{S}^n)$ (the n th de Rham cohomology group of the sphere \mathbb{S}^n). Note that the generalized Stokes integral theorem tells us that

$$\int_{\mathbb{S}^n} d\nu = \int_{\partial\mathbb{S}^n} \nu = 0,$$

since the boundary $\partial\mathbb{S}^n$ of the sphere \mathbb{S}^n is empty (see page 729). Therefore,

$$\chi(\mathbb{S}^n) = \int_{\mathbb{S}^n} \omega + d\nu,$$

that is, the Euler characteristic of the sphere only depends on the Euler class of the sphere. Let $n = 1, 2, 3, \dots$. A special case of the Poincaré–Hopf theorem tells us that:

A continuous velocity vector field without stagnation points exists on the n -dimensional sphere iff $\chi(\mathbb{S}^n) = 0$.

Explicitly, this means the following:

- If n is even, then every continuous velocity vector field on \mathbb{S}^n has a stagnation point.
- If n is odd, then there exists a continuous velocity vector field on \mathbb{S}^n which has no stagnation point.

The theorem of Adams. By definition, an m -field on the sphere \mathbb{S}^n is a family of m continuous velocity vector fields $\mathbf{v}_1, \dots, \mathbf{v}_m$ on \mathbb{S}^n with the property that the tangent vectors $\mathbf{v}_1(P), \dots, \mathbf{v}_m(P)$ are linearly independent at each point P in \mathbb{S}^n . The number

$$\text{Span}(T\mathbb{S}^n)$$

is the maximal number m such that an m -field exists on \mathbb{S}^n . We have:

- $\text{Span}(T\mathbb{S}^n) = 0$ if n is even (i.e., $n = 2, 4, 6, \dots$),
- $\text{Span}(T\mathbb{S}^1) = 1$, $\text{Span}(T\mathbb{S}^3) = 3$, $\text{Span}(T\mathbb{S}^5) = 1$, $\text{Span}(T\mathbb{S}^7) = 7$.

The deep theorem of Adams tells us the precise result:⁴⁴

$$\text{Span}(T\mathbb{S}^n) = 8a + 2^b - 1, \quad n = 1, 2, \dots$$

Here, the nonnegative integers a and b with $b \leq 3$ are uniquely determined by the prime number factorization $n + 1 = 2^{4a+b} \cdot c$, where c is a positive odd integer.

For example, if $n = 7$, then $8 = 2^3 \cdot 1$. Hence $a = 0, b = 3, c = 1$. This implies $\text{Span}(T\mathbb{S}^7) = 2^3 - 1 = 7$.

Global parallel transport. The sphere \mathbb{S}^n is called parallelizable iff the condition $\text{Span}(T\mathbb{S}^n) = n$ is satisfied.

The sphere \mathbb{S}^n is parallelizable iff $n = 1, 3, 7$.

Let us discuss this. Suppose that $n = 1, 3, 7$. We want to show that there exists a global parallel transport. In fact, there exist continuous velocity fields $\mathbf{v}_1, \dots, \mathbf{v}_n$ on \mathbb{S}^n such that the vectors

$$\mathbf{v}_1(P), \dots, \mathbf{v}_n(P)$$

form a basis of the tangent space $T_P\mathbb{S}^n$ for all points $P \in \mathbb{S}^n$. Fix the point P_0 and choose a fixed tangent vector $\mathbf{v}_0 \in T_{P_0}$. Then there exist uniquely determined real numbers v^1, \dots, v^n such that

$$\mathbf{v}_0 = \sum_{j=1}^n v^j \mathbf{v}_j(P_0).$$

Naturally enough, the global parallel transport of the vector \mathbf{v}_0 to the arbitrary point P of \mathbb{S}^n is defined by

$$\mathbf{v}(P) := \sum_{j=1}^n v^j \mathbf{v}_j(P).$$

Triviality of the tangent bundle. The tangent bundle $T\mathbb{S}^n$ is called trivial iff \mathbb{S}^n is parallelizable, that is, $n = 1, 3, 7$. Then every tangent vector \mathbf{v} at the point P can uniquely be represented by the formula

$$\mathbf{v} = \sum_{j=1}^n v^j \mathbf{v}_j(P)$$

⁴⁴ J. Adams, Vector fields on spheres, Ann. Math. **75** (1962), 603–632.

where v^1, \dots, v^n are arbitrary real numbers. This means that we have a global coordinate system for the tangent bundle TS^n . There exists a bijection between the points (P, \mathbf{v}) of the tangent bundle and the points

$$(P, v^1, \dots, v^n)$$

of the product bundle $S^n \times \mathbb{R}^n$. Summarizing:

The topology of fiber bundles allows us to distinguish between trivial and nontrivial physical fields.

We expect that essential physical effects are related to nontrivial topological structures.

Typical examples. In the present Volume III, we will study the following applications of gauge theory:

- gauge theory on the 3-dimensional Euclidean manifold (Sect. 9.4),
- gauge theory on the sphere (e.g., the surface of earth) as a paradigm (Sect. 9.5),
- the relation between Gauss' surface theory and Levi-Civita's parallel transport (Sect. 9.5),
- the relation between Gauss' surface theory and Cartan's method of moving frames (Sect. 9.5),
- the main theorem on velocity vector fields on the Euclidean manifold – the classic predecessor of gauge theory in physics (Sect. 12.10.3),
- Maxwell's theory of electromagnetism as a commutative $U(1)$ -gauge theory on the Minkowski manifold (Chap. 13 and Chaps. 18–23),
- the noncommutative $SU(N)$ -gauge theory due to Yang and Mills (Chap. 15).

The general axiomatic approach to gauge theory will be discussed in Chap. 17. In Volume IV, we will study

- the Standard Model in particle physics with a representation of the product group $U(1) \times SU(2) \times SU(3)$ as gauge group, and
- the general theory of relativity for gravitation (connection on the tangent bundle of the pseudo-Riemannian space-time manifold),
- minimal surfaces and the conformal gauge symmetry,
- string theory and the conformal gauge symmetry.

Perspectives

The discussion above displays fruitful relations between mathematics and physics. Let us add some further quotations.

Instantons and gauge theory:

From 1977 onward my interest moved in the direction of gauge theories and the interactions between geometry and physics. I had for many years a mild interest in theoretical physics, stimulated on many occasions by lengthy discussions with George Mackey from Harvard University. However, the stimulus in 1977 came from two other sources. On the one hand, Singer told me about the Yang–Mills equations, which through the influence of Yang were just beginning to percolate into mathematical circles. During his stay in Oxford in early 1977, Singer, Hitchin, and I took a serious look at the self-duality equations. We found that a simple application of the Atiyah–Singer index theorem gave the formula for the number of instanton parameters ... The other stimulus came from the presence in Oxford of

Roger Penrose and his group working on relativistic spinor calculus and twistor theory.⁴⁵

Sir Michael Atiyah, 1988

Conformal symmetry and twistors:

A new type of algebra for Minkowski space-time is described, in terms of which it is possible to express any conformally or Poincaré covariant operation. The elements of the algebra (twistors) are combined according to tensor-type rules, but they differ from tensors or spinors in that they describe locational properties in addition to directional ones.

Twistor algebra will have the same type of universality, in relation to the conformal group, that the well-known and highly effective two-component spinor algebra of van der Waerden has, in relation to the Lorentz group. Twistors are, in fact, the “spinors” which are relevant to the six-dimensional space whose (pseudo)-rotation group is isomorphic to the conformal group of ordinary space-time.⁴⁶

Roger Penrose, 1967

The Seiberg–Witten equations and the quark confinement:

Riemannian, symplectic and complex geometry are often studied by means of solutions to systems of nonlinear differential equations, such as the equations of geodesics, minimal surfaces, Einstein’s curved universe, pseudo-holomorphic curves and Yang–Mills connections. For studying such equations, a unified technology has been developed, involving analysis on infinite-dimensional manifolds.

A striking application of the new technology is Donaldson’s theory of “anti-self-dual” connections on $SU(2)$ -bundles over four-manifolds, which applies the Yang–Mills equations from mathematical physics to shed light on the relationship between the classification of topological and smooth four-manifolds. This reverses the expected direction of application from topology to differential equations to mathematical physics. Even though the Yang–Mills equations are only mildly nonlinear, a prodigious amount of nonlinear analysis is necessary to fully understand the properties of the space of solutions.

At our present state of knowledge, understanding smooth structures on topological four-manifolds seems to require nonlinear as opposed to linear partial differential equations. It is therefore quite surprising that there is a set of partial differential equations which are even less nonlinear than the Yang–Mills equation, but can yield many of the most important results from Donaldson’s theory. These are the Seiberg–Witten equations . . .

During the 1980’s, Simon Donaldson used the Yang–Mills equations, which had originated in mathematical physics, to study the differential topology

⁴⁵ See M. Atiyah, *Collected Works*, vol. V: *Gauge Theories*, Clarendon Press, Oxford, 1988. Reprinted by permission of Oxford University Press.

⁴⁶ Reprinted with permission from R. Penrose, *Twistor algebra*, *J. Math. Phys.* **8**(2) (1967), 345–366. Copyright 1967, American Institute of Physics.
N. Hitchin, G. Segal, and R. Ward, *Integrable Systems, Twistors, Loop Groups, and Riemann Surfaces*, Oxford University Press, 1999.
J. Frauendiener and R. Penrose, *Twistors and general relativity*, pp. 479–505. In: B. Engquist and W. Schmid (Eds.), *Mathematics Unlimited—2001 and Beyond*, Springer, New York, 2001.

of four-manifolds. Using work of Michael Freedman, he was able to prove theorems of the following type:

- There exist many compact four-manifolds which have no smooth structure.
- There exist many pairs of compact four-manifolds which are homeomorphic but not diffeomorphic.

The nonlinearity of the Yang-Mills equations presented difficulties, so many techniques within the theory of nonlinear partial differential equations had to be developed. Donaldson's theory was elegant and beautiful, but the details were difficult for beginning students to master.

In the fall of 1994, the physicist Edward Witten proposed a set of equations which give the main results of Donaldson's theory in a far simpler way than had been thought possible... The Seiberg–Witten equations give rise to new invariants of four-dimensional smooth manifolds called the Seiberg–Witten invariants. The key point is that homeomorphic smooth four-manifolds may have quite different Seiberg–Witten invariants... Shortly after the Seiberg–Witten invariants were discovered, several striking applications were found concerning

- (i) the proof of the Thom conjecture on the smooth embedding of compact Riemann surfaces into two-dimensional complex projective spaces \mathbb{P}_C^2 ,
- (ii) obstructions to the existence of a Riemannian geometry with positive curvature on manifolds, and
- (iii) the existence of pseudo-holomorphic curves on symplectic manifolds.⁴⁷

John Moore, 1996

This quotation refers to the following two fundamental papers on the quark confinement:

N. Seiberg and E. Witten, Electric-magnetic duality, monopole condensation, and confinement in $N = 2$ supersymmetric Yang–Mills theory, *Nuclear Phys.* **B426** (1994), 19–52.

N. Seiberg and E. Witten, Monopoles, duality and chiral symmetry breaking in $N = 2$ supersymmetric QCD, *Nucl. Physics* **B431** (1994), 485–550.

These two papers concern the computation of models which describe electrically and magnetically charged supersymmetric particles at low energies in the setting of gauge theory. The Seiberg–Witten equations use the spin structure of manifolds called spin manifolds, and they generalize the Landau–Ginzburg equation in superconductivity. This can be found in:

C. Nash, *Topology and Physics – a Historical Essay*, pp. 359–415. In: I. James (Ed.), *History of Topology*, Oxford University Press, 1999.

M. Atiyah, *The Dirac equation and geometry*, pp. 108–124. In: P. Goddard (Ed.), *Paul Dirac – the Man and his Work*, Cambridge University Press, 1998.

E. Witten, *Physical law and the quest for mathematical understanding*, *Bull. Amer. Math. Soc.* **40** (2003), 21–30.

⁴⁷ J. Moore, *Lectures on Seiberg–Witten Invariants*, Springer, Berlin, 1996 (reprinted with permission). We recommend these Lecture Notes as an introduction to the study of the Seiberg–Witten equations.

E. Witten, From superconductors and four-manifolds to weak interaction, *Bull. Math. Soc.* **44** (2007), 361–391.

J. Jost, *Riemannian Geometry and Geometric Analysis*, Springer, Berlin, 2008.

H. Lawson and M. Michelsohn, *Spin Geometry*, Princeton University Press, 1994.

Furthermore, we refer to:

R. Stern, Instantons and the topology of four-manifolds, *Mathem. Intelligencer* **5**(3) (1983), 39–44.

K. Marathe, *Topics in Physical Mathematics*, Springer, London, 2010.

D. Freed and K. Uhlenbeck, *Instantons and Four-Manifolds*, Springer, New York, 1984.

S. Donaldson and P. Kronheimer, *The Geometry of Four-Manifolds*, Oxford University Press, 1990.

S. Donaldson, The Seiberg–Witten equations and 4-manifold topology, *Bull. Amer. Math. Soc.* **33** (1996), 45–70.

J. Morgan, *The Seiberg–Witten Equations and Applications to the Topology of Four-Manifolds*, Princeton University Press, 1996.

T. Friedrich, *Dirac Operators in Riemannian Geometry*, Amer. Math. Soc., Providence, Rhode Island, 2000.

P. Kronheimer and T. Mrowka, *Monopoles and Three-Manifolds*, Cambridge University Press, 2007.

Concerning Morse theory, Floer homology, and quantum cohomology, we refer to the following monographs:

J. Jost, *Riemannian Geometry and Geometric Analysis*, Springer, Berlin, 2008.

M. Schwarz, *Morse Homology*, Birkhäuser, Basel, 1993.

D. McDuff and D. Salamon, *J-holomorphic Curves and Quantum Cohomology*, Amer. Math. Soc., Providence, Rhode Island, 1994.

Yu. Manin, *Frobenius manifolds, quantum cohomology and moduli spaces*, Amer. Math. Soc., Providence, Rhode Island, 1999.

S. Donaldson, *Floer Homology Groups*, Cambridge University Press, 2002.

J. Kock and I. Vainsencher, *An Invitation to Quantum Cohomology: Kontsevich's Formula for Plane Curves*, Birkhäuser, Basel, 2006.

S. Novikov, *Topological Library*, Vol. 1: Cobordisms and their Applications, Vol. 2: Characteristic Classes and Smooth Structures, World Scientific, Singapore, 2007/09.

The spectrum of elliptic Dirac operators on manifolds and noncommutative geometry. The original Dirac equation for the relativistic electron is a first-order system of partial differential equations of hyperbolic type. In geometry, one uses an elliptic variant of the Dirac operator by passing to imaginary time. For example, the Seiberg–Witten equations in geometry are nonlinear equations related to the elliptic Dirac differential operator. In 1985, Alain Connes created noncommutative geometry. The decisive analytic information comes from the spectrum of an elliptic Dirac operator on a compact manifold. References can be found on page 346.

The Millennium Prize Problem in quantum field theory. One of the seven Millennium Prize Problems concerns the Yang–Mills gauge theory. This is described in:

A. Jaffe and E. Witten, Quantum Yang–Mills theory, pp. 129–152. In: J. Carlson, A. Jaffe, and A. Wiles (Eds.), *The Millennium Prize Problems*, Amer. Math. Soc., Providence, Rhode Island, 2006.

The problem is to show that there exists a gap between the ground state energy and the first excitation energy of a Yang–Mills quantum field. The award for solving this problem will be one million dollars. Nowadays it is completely open how to attack this problem.

The Seiberg–Witten equations and the Weinstein conjecture. Let us mention a recent beautiful and deep application of the Seiberg–Witten equations to dynamical systems on 3-dimensional contact manifolds. Recall first the classical Poincaré theorem saying that every continuous velocity vector field on a two-dimensional sphere has a zero. In terms of physics, this corresponds to a stationary point of the velocity vector field. Such a point represents a (trivial) closed orbit of the corresponding flow of fluid particles. In other words, the Poincaré theorem tells us that:

Every continuous velocity vector field on a 2-dimensional sphere has a closed orbit.

We want to generalize this to the 3-dimensional sphere. It turns out that this is a hard problem. First of all note that there exist continuous velocity vector fields on the 3-dimensional sphere which have no closed orbits. Recently, Taubes proved the following theorem:

Every smooth Reeb velocity vector field on a real compact oriented 3-dimensional manifold (without boundary) has a closed orbit.

This tells us that the Weinstein conjecture is true in three dimensions. Let us explain the notation. A Reeb velocity vector field \mathbf{v} on the 3-dimensional manifold \mathcal{M} is given by the equations

$$d\omega(\mathbf{v}) = 0, \quad \omega(\mathbf{v}) = 1 \quad \text{on } \mathcal{M}$$

where ω is a contact form on \mathcal{M} , that is, ω is a smooth 1-form on \mathcal{M} with the property $(d\omega \wedge \omega)(P) \neq 0$ for all points P of \mathcal{M} .⁴⁸

The point of departure is the Taubes theorem on the relation between the Seiberg–Witten theory and the Gromov theory on pseudo-holomorphic curves. This theorem relates the Seiberg–Witten invariants of real symplectic 4-dimensional manifolds to counts of holomorphic curves. In the present case, we have to deal with 3-dimensional manifolds. To this end, Taubes uses a 3-dimensional variant of the Seiberg–Witten theory combined with Floer homology. We refer to:

M. Hutchings, Taubes’s proof of the Weinstein conjecture in dimension three, *Bull. Amer. Math. Soc.* **47**(1) (2010), 73–126.

C. Taubes, *Seiberg–Witten and Gromov Invariants for Symplectic 4-Manifolds*, International Press, Boston, 2000.

C. Taubes, The Seiberg–Witten equations and the Weinstein conjecture I, II: *Geom. Topol.* **11** (2007), 2117–2002, **13** (2009), 1337–1417.

Taubes discovered in the 1990s that

⁴⁸ An introduction to Lie’s contact geometry can be found in Sect. 5.7 of Vol. II.

The Seiberg–Witten theory and Gromov’s theory of pseudo-holomorphic curves are equivalent in some sense.

This relates apparently completely different physical and mathematical topics with each other.

- In terms of physics, the Seiberg–Witten equation is closely related to the Landau–Ginzburg equation for describing phase transitions in condensed matter (e.g., superconductors).
- The Gromov theory of pseudo-holomorphic curves generalizes the Cauchy–Riemann differential equations for holomorphic functions.

We refer to:

I. Vekua, Generalized Analytic Functions, Pergamon Press, London, 1962.

M. Gromov, Pseudo-holomorphic curves in symplectic manifolds, Invent. Math. **82** (1985), 307–347.

H. Hofer and E. Zehnder, Symplectic Invariants and Hamiltonian Dynamics, Birkhäuser, Basel, 1994.

D. McDuff and D. Salamon, *J*-holomorphic Curves and Quantum Cohomology, Amer. Math. Soc., Providence, Rhode Island, 1994.

D. McDuff and D. Salamon, Introduction to Symplectic Topology, Clarendon Press, Oxford, 1998.

We will show in Sect. 2.3.2 on page 129 that the Cauchy–Riemann equations are closely related to the Clifford algebra $\mathbb{V}(E_2)$ of the Euclidean plane E_2 (i.e., the algebra of quaternions). It turns out that

The Dirac equation for the relativistic electron is based on the Clifford algebra $\mathbb{V}(M_4)$ of the 4-dimensional Minkowski space M_4 (flat space-time manifold). Based on the theory of Clifford algebras, the Dirac equation represents a generalization of the classical Cauchy–Riemann equations.

In his 1851 Ph.D. thesis, Riemann used the Cauchy–Riemann differential equations in order to create the geometric theory of holomorphic functions based on the notion of conformal map and the idea of the Riemann surface for describing analytic continuation in a global setting. Riemann was strongly motivated by ideas coming from physics (e.g., electricity). He used physical intuition in order to motivate the existence of global analytic functions on compact Riemann surfaces. Riemann’s successors filled the gaps in Riemann’s arguments step by step. The final form of the theory was published by

H. Weyl, The Concept of a Riemann Surface (in German), Teubner, Leipzig, 1913. New edition with commentaries supervised by R. Remmert, Teubner, Leipzig, 1997. English edition: Addison Wesley, Reading, Massachusetts, 1955.

In terms of mathematics, the following topics lurk behind the equivalence of the Seiberg–Witten theory and the Gromov theory of pseudo-holomorphic curves:

- symplectic geometry, fixed-point theorems for symplectic maps (i.e., higher-dimensional versions of Poincaré’s last theorem), geometrical optics, and Hamiltonian mechanics (e.g., celestial mechanics) (see Vol. II),
- Lie’s contact geometry and the Legendre transformation for thermodynamical potentials (see Vol. II),
- Clifford algebras, spin geometry, the Dirac equation for the relativistic electron in flat and curved space-times, symmetry breaking and the Higgs particle in weak interaction, and the quark confinement.

The creation of Floer homology was essentially motivated by the following fundamental paper:

E. Witten, Supersymmetry and Morse theory, J. Diff. Geom. **17** (1982), 661–692.

We refer to the basic paper by

A. Floer, Witten's complex and infinite-dimensional Morse theory, J. Diff. Geometry **30** (1989), 207–221.

Morse theory studies the relation between the critical points of energy functionals $E : \mathcal{M} \rightarrow \mathbb{R}$ and the topology of the manifold \mathcal{M} . As a nice introduction to modern Morse theory based on Floer homology, we recommend:

J. Jost, Riemannian Geometry and Geometric Analysis, fifth edition, Chapter 6, Springer, Berlin, 2008.

We also recommend the monographs on Morse homology and its applications by Schwarz (1993) and Donaldson (2002) quoted on page 60.

The language of bundles. The following equations of physical theories possess a similar structure:

- the basic equations of quantum electrodynamics (see Chap. 11 of Vol. II),
- the Landau–Ginzburg equation,
- the Dirac equation for the relativistic electron,
- the Yang–Mills equation,
- the Seiberg–Witten equation,
- the Standard model in particle physics (see Vol. IV).

In order to display the similarities, one has to use variational problems based on the principle of critical action. Then the Lagrangians possess a similar structure which depends on the choice of both

- the symmetry group \mathcal{G} (the curvature of the principal fiber bundle \mathcal{P} with the structure group \mathcal{G}), and
- the spaces of physical fields (sections of vector bundles which are associated to \mathcal{P} via representations of the symmetry group \mathcal{G}).

The symmetry group of quantum electrodynamics (resp. of the Yang–Mills equations) is the group $U(1)$ (resp. $SU(2)$). The Dirac equation and the Seiberg–Witten equations are based on so-called spin groups (universal covering groups of the Lorentz group $O(1, 3)$ and the rotation groups $SO(N)$) which are closely related to the spin of the particles.

The interactions correspond to the curvature of the principal fiber bundle.

The basic ideas will be studied in Chap. 15 (Ariadne's thread in gauge theory).

Gauge Potentials, moduli spaces, correlation functions of quantum fields, and Feynman path integrals. The basic formula

$$F = DA$$

describes the relation between the interaction forces F and the gauge potential A by means of the first-order differential operator D . Note the following peculiarity. By a local gauge transformation, we understand a change of the local bundle coordinates of the corresponding principal bundle with the symmetry group \mathcal{G} . By a global gauge transformation, roughly speaking, we understand a diffeomorphism $f : \mathcal{P} \rightarrow \mathcal{P}$ of the bundle space \mathcal{P} of the principal bundle which is generated by the action of the symmetry group \mathcal{G} (also called gauge group) on the principal fiber bundle. There arises the following question:

Which is the qualitative (topological) and quantitative structure of the space of all gauge potentials A up to global gauge transformations?

In other words, one has to investigate the moduli space $\text{Mod}(A)$ of all gauge potentials A (called connections in mathematics). More precisely, $\text{Mod}(A)$ is the space of all equivalence classes of connections modulo global gauge transformations. The space $\text{Mod}(A)$ is called the moduli space of gauge potentials (connections). The investigation of moduli spaces needs sophisticated topological tools. The point is that, as a rule, moduli spaces are not smooth structures; they are not manifolds, but they possess singularities. In a natural way, such geometric objects arise in algebraic geometry (e.g., the curve $x^2 - y^2 = 0$ has a singularity at the point $(0, 0)$). We refer to:

S. Abhyankar, *Algebraic Geometry for Scientists and Engineers*, Amer. Math. Soc., Providence, Rhode Island, 1990.

In terms of quantum field theory, the knowledge of the moduli space is fundamental for computing correlation functions via Feynman path integrals:

$$\int_{[A] \in \text{Mod}(A)} e^{iS([A])/\hbar} \mathcal{D}([A]). \quad (0.92)$$

Here, one has to sum (i.e., to integrate) over all the possible physical states $[A]$ (i.e., over all the elements of the moduli space). The statistical weight $e^{iS([A])/\hbar}$ depends on the action $S([A])$ corresponding to the physical state $[A]$ (equivalence class of connections modulo global gauge transformations). As an introduction to moduli spaces in gauge theory, we recommend:

K. Marathe and G. Martucci, *The Mathematical Foundations of Gauge Theories*, North-Holland, Amsterdam, 1992.

K. Marathe, *Topics in Physical Mathematics*, Springer, London, 2010.

G. Naber, *Topology, Geometry, and Gauge Fields*, Springer, New York, 1997.

Fundamental papers on this subject can be found in:

M. Atiyah, *Collected Works*, Vol. 5: *Gauge Theories*, Cambridge University Press, 2004.

Topological quantum field theory. The basic idea of topological quantum field theory is to choose special functionals $A \mapsto S([A])$ in order to obtain topological invariants by using Feynman path integrals of the type (0.92). These integrals are computed by means of the method of stationary phase. This is an approximative method. However, there exists a rigorous result which shows that, in a special model, the method of stationary phase in lowest order yields the precise value of the integral:

J. Duistermaat and G. Heckmann, On the variation in the cohomology in the symplectic form of the reduced phase space, *Invent. Math.* **69** (1982), 259–268; **72** (1983), 153.

We refer to:

E. Witten, Topological quantum field theory, *Commun. Math. Phys.* **117** (1988), 353–386.

E. Witten, *Witten's Lectures on Three-Dimensional Topological Quantum Field Theory*. Edited by Sen Hu, World Scientific, Singapore 1999.

In Sect. 23.8, we will sketch how the Jones polynomials (i.e., topological invariants in knot theory) can be obtained by the method of topological quantum field theory due to Witten. This approach is based on the Chern–Simons gauge theory on the 3-dimensional sphere.

Historical remarks on moduli spaces and modular forms. The moduli space $\text{Mod}_g(\mathbf{R})$ of compact Riemann surfaces \mathbf{R} of genus g consists of all equivalence classes of compact Riemann surfaces of genus g modulo conformal equivalence. This space was first studied by Riemann who determined the finite dimension of this space:

- If $g \geq 2$, then the real dimension of $\text{Mod}(\mathbf{R})_g$ is equal to $6g - 6$. This corresponds to the conformal classification of algebraic curves parametrized by sophisticated automorphic functions which were investigated by Poincaré and Klein at the end of the 19th century.
- If $g = 0$, then the moduli space $\text{Mod}(\mathbf{R})_0$ consists of precisely one point which corresponds to the Riemann sphere; this sphere is conformally equivalent to the one-dimensional complex projective space $\mathbb{P}_{\mathbb{C}}^1$.
- If $g = 1$, then the dimension of $\text{Mod}(\mathbf{R})_1$ is equal to two. This corresponds to the conformal classification of elliptic curves parametrized by elliptic functions. The theory of elliptic functions was created by Legendre, Gauss, Jacobi, and Weierstrass at the end of the 18th century and in the 19th century.

In the setting of Teichmüller theory, the moduli space $\text{Mod}(\mathbf{R})_g$ is studied in:

J. Jost, *Compact Riemann Surfaces: An Introduction to Contemporary Mathematics*, Springer, Berlin, 2006.

Furthermore, we refer to:

F. Klein, *Development of Mathematics in the 19th Century*, Math. Sci. Press, New York, 1979.

K. Maurin, *The Riemann Legacy: Riemannian Ideas in Mathematics and Physics of the 20th Century*, Kluwer, Dordrecht, 1997.

A. Hurewicz and R. Courant, *Lectures on Complex Function Theory and Elliptic Integrals* (in German), Springer, Berlin, 1964.

D. Lawden, *Elliptic Functions and Applications*, Springer, New York, 1989.

L. Ford, *Automorphic Functions*, Chelsea, New York, 1972.

T. Apostol, *Introduction to Analytic Number Theory*, Springer, New York, 1986.

T. Apostol, *Modular Functions and Dirichlet Series in Number Theory*, Springer, New York, 1990.

M. Waldschmidt, P. Moussa, J. Luck, and C. Itzykson (Eds.), *From Number Theory to Physics*, Springer, New York, 1995 (survey articles).

J. Bruinier, G. van der Geer, G. Harder, and D. Zagier, *The 1-2-3 of Modular Forms, Lectures at a Summer School in Nordfjordeid, Norway*, Springer, Berlin, 2008 (survey articles).

The theory of elliptic curves possesses a very rich structure. The spectacular proof of Fermat's last theorem by Andrew Wiles in 1995 was based on recent progress for elliptic curves (see the discussion in the Prologue to Vol. I on page 17):

S. Singh, *Fermat's Last Theorem: The Story of a Riddle that Confounded the World's Greatest Minds for 358 Years*, Fourth Estate, London, 1997.

F. Diamond and J. Shurman, *A First Course in Modular Forms*, Springer, Berlin, 2005 (the modularity theorem).

Fermat's last theorem is a consequence of the so-called modularity theorem:

All rational elliptic curves arise from modular forms.

This fundamental result was conjectured by Taniyama, Shimura, and Weil in the 1950s and 1960s. For a special class of elliptic curves, the theorem was proven by Wiles in order to get the proof of Fermat's last theorem. The correctness of the general modularity theorem was proven by Breuil, Conrad, Diamond, and Taylor in 2001:

A. Wiles, Modular elliptic curves and Fermat's last theorem, *Ann. Math.* **142** (1995), 443–551.

C. Breuil, B. Conrad, F. Diamond, and R. Taylor, On the modularity of elliptic curves over the field \mathbb{Q} of rational numbers: wild 3-adic exercises, *J. Amer. Math. Soc.* **14**(4) (2001), 843–939.

An introduction to the field \mathbb{Q}_p of p -adic numbers and the adelic ring $\mathbb{A}_{\mathbb{Q}}$ will be given in Sect. 4.6.7 on page 332. We will also briefly discuss the relation of p -adic numbers and the adelic ring to mathematical models motivated by physics (chaos and turbulence or dark matter in cosmology). This is called adelic physics.

The moduli space of compact Riemann surfaces is used in string theory in order to compute the Feynman path integral for the free bosonic string. This was first done by:

A. Polyakov, Quantum geometry of bosonic strings, *Phys. Lett.* **103B** (1981), 207.

A. Polyakov, Quantum geometry of fermionic strings, *Phys. Lett.* **103B** (1981), 213.

A detailed computation can be found in:

B. Hatfield Quantum Field Theory of Point Particles and Strings, Addison-Wesley, Redwood City, California.

For a rigorous approach based on the mathematical theory of Riemann surfaces, we refer to:

J. Jost, The Bosonic String: A Mathematical Treatment, International Press, Boston, 2001.

The Ricci flow and the Poincaré conjecture:

The system of ordinary differential equations

$$\boxed{\frac{\partial \mathbf{g}}{\partial \sigma} = -\text{Ric}(\mathbf{g})} \quad (0.93)$$

defines the Ricci flow $\mathbf{g} = \mathbf{g}(P, \sigma)$, $P \in \mathcal{M}$, on the n -dimensional (compact) Riemannian manifold \mathcal{M} . This is a family of metric tensors on \mathcal{M} which depends on the real parameter $\sigma \in [\sigma_0, \sigma_1]$. In terms of local coordinates, equation (0.93) reads as

$$\frac{\partial g_{ij}(x, \sigma)}{\partial \sigma} = -R_{ij}(g(x, \sigma)), \quad x \in \mathcal{M}, \sigma \in [\sigma_0, \sigma_1]$$

where $i, j = 1 \dots, n$. This equation generalizes the heat equation

$$\frac{\partial \Theta(P, \sigma)}{\partial \sigma} = -\Delta \Theta(P, \sigma), \quad P \in \mathcal{M}, \sigma \in [\sigma_0, \sigma_1]$$

for the temperature field $\Theta = \Theta(P, \sigma)$ on the manifold \mathcal{M} . Here, the real parameter σ denotes time. It was the ingenious idea of Perelman to solve the Poincaré conjecture by deforming appropriate 3-dimensional manifolds to a 3-dimensional sphere by means of the Ricci flow. We refer to:

B. Chow and D. Knopf, *The Ricci Flow: An Introduction*, Amer. Math. Soc., Providence, Rhode Island, 2004.

B. Chow, P. Lu, and L. Ni, *Hamilton's Ricci Flow*, Amer. Math. Soc., Providence, Rhode Island, 2006.

H. Cao, S. Yau, and X. Zhu, *Structure of Three-dimensional Space: The Poincaré and Geometrization Conjectures*, International Press, Boston, 2006.

J. Morgan and G. Tian, *Ricci Flow and the Poincaré Conjecture*, Amer. Math. Soc., Providence, Rhode Island/Clay Mathematics Institute, Cambridge, Massachusetts, 2007.

D. O'Shea, *The Poincaré Conjecture: In Search of the Shape of the Universe*, Walker, New York, 2007.

Modern differential geometry in physics. A lot of material on modern differential geometry and its applications to physics can be found in:

V. Ivancevic and T. Ivancevic, *Differential Geometry: A Modern Introduction*, World Scientific, Singapore, 2007.

A lot of historical material including the history of gauge field theory is contained in the survey article by:

H. Kastrup, On the advancement of conformal transformations and their associated symmetries in geometry and theoretical physics, *Annalen der Physik* **17** (2008), 631–690.

For the relations between mathematics and physics in the history of quantum field theory, we recommend:

W. Nahm, Conformal field theory: a bridge over troubled waters, pp. 571–604. In: A. Mitra (Ed.), *Quantum Field Theory: A 20th Century Profile*, Indian National Science Academy and Hindustan Book Agency, India, 2000.

For recent progress in quantum field theory based on close relations between mathematics and physics, we refer to:

C. Bär and K. Fredenhagen, *Quantum Field Theory in Curved Space-Times*, Springer 2009.

<http://www.springer.com/978-3-642-22420-1>

Quantum Field Theory III: Gauge Theory
A Bridge between Mathematicians and Physicists
Zeidler, E.
2011, XXXII, 1126 p., Hardcover
ISBN: 978-3-642-22420-1