B

A lexicon

When describing previously uncharted territories, discoverers and inventors are forced to adopt and adapt previously known terms, concepts and techniques for the new phenomena, or invent wholly new ones. This appendix collects a listing of perhaps less familiar but oft-used terms in our field, then turns to the vector/tensor and even functional extension to the hopefully well-familiar rules of multivariate calculus, and closes with a brief on Gödel's incompleteness theorem.

B.1 The jargon

The jargon of theoretical and mathematical physics is very much in development and in some cases not yet standardized. With the aim of using compact but precise terms to name very specific ideas, many scientists begin using an otherwise rarely used word and, at times, their choice "catches on" and becomes standardized. At other times, different terms are used by competing (or non-communicating) research groups for the same or closely related concepts, whereupon one of the two "competing" terms may turn into a standard but only after a long period during which both terms are used. As the fundamental physics of elementary particles is still very much in development, consistency and expediency required me to make certain choices in terminology, which I have, to the best of my knowledge, indicated together with possible alternatives.

The subsequent lexicon offers brief explanations for some of the perhaps less familiar technical terms and expressions, most of which are fairly standard, but in a field other than particle physics.

Abelian (commutative, symmetric) A binary operation $\star$ is abelian if $a \star b = b \star a$. By extension, structures defined using an abelian binary operation are also called abelian. Operations that are not abelian are called non-abelian (= non-commutative, = asymmetric), as are structures defined using them.

Algebra A vector space $\mathfrak{A}$ over a field $\mathbb{k}$, equipped with a binary operation $*$, which satisfies the distribution law over addition: $a * (b + c) = (a * b) + (a * c)$, for all elements $a, b, c \subset \mathfrak{A}$, and for which it is true that $\alpha(a * b) = (\alpha a) * b = a * (\alpha b)$, for each $\alpha \in \mathbb{k}$ and $a, b \in \mathfrak{A}$. The operation $*$ is typically a type of multiplication; it is often commutative, i.e., symmetric, but in Lie algebras it is antisymmetric: $a * b = -b * a$.

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

Amplitude In the context of field theory and so also in (high energy) elementary particle physics, this is the matrix element $\mathfrak{M}_{i\to f} := \langle f|H_{\rm int}|i\rangle$ where $|i\rangle$ and $|f\rangle$ are the initial and final states and $H_{\rm int}$ is the (algebraic sum of all) interaction operator(s) that can bring about the process $|i\rangle \to |f\rangle$. The probability for this process is then proportional to $|\mathfrak{M}_{i\to f}|^2$ [☞ display (3.85) and Section 3.3.3 on p. 113].

Analytic function A function $f(x)$ is analytic in a domain $\mathscr{D}$ if it has a (convergent) Taylor expansion $f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n$ for every $x_0, (x-x_0) \in \mathscr{D}$.

Anomaly Structural changes in relations between observables caused by passing from classical to quantum theory. If those relations represent the algebra of symmetry transformations and anomalies obstruct the closure or change the structure of that algebra, then anomalies destroy or change the symmetry that was built into the system originally – which points to an inconsistency. Models with an anomaly in a gauge symmetry are simply inconsistent [☞ Section 7.2.3], whereupon all gauge anomaly ought to cancel. [☛ geometric quantization; canonical quantization] In turn, anomalies in global and approximate symmetries need not cancel, but are characteristic quantities that cannot be altered by field redefinitions, and so must remain conserved throughout the evolution of a system, including phase transitions. This is a direct consequence of the underlying principle in Dirac quantization. [☛ Dirac quantization]

Auxiliary field A field that has a non-differential equation of motion, which determines the field point-by-point. If this equation of motion can be solved, the solution can be reinserted in the Lagrangian density, which is classically equivalent to the original Lagrangian density but involves fewer fields. The equivalence need not hold between the quantum models defined from the two Lagrangian densities.

Baryon Since the acceptance of the quark model in 1973, a bound state of three quarks. Originally, a particle that interacts by means of the strong nuclear force (at $\sim 10^{-23}$ s), can be detected as an isolated particle, and has a mass that is not smaller than that of the proton, such as a neutron.

Bijection A mapping $f : X \to Y$ that is both (**1**) an ***injection*** (i.e., “1–1”), so for every $x \in X$ there is precisely one $y = f(x) \in Y$, and (**2**) a ***surjection***, so for every $y \in Y$ there is an $x \in X$ so that $f(x) = y$. Bijection = surjective injection, i.e., injective surjection. [☛ injection, surjection]

BFV-quantization A contemporary version (by Igor Batalin, Efim S. Fradkin and Grigori Vilkovisky) of canonical quantization in the Hamiltonian formalism, which generalizes the evolution of the canonical–Dirac–BRST quantization to the general case when the constraints do not close the structure of an algebra [174, 39, 172, 36, 37, 345, 38, and references therein]; see also the texts [268, 555, 484, 496, 589, 590] and [509]. [☛ BRST quantization; Dirac quantization; canonical quantization]

Bose condensation The state of a system where infinitely many particles (bosons) are in the same quantum state. The Coulomb static potential may be understood as a Bose condensation of infinitely many photons.

Boson By definition, a particle (as well as its mathematical representatives: wave-functions, creation and annihilation operators or fields) that obeys the Bose–Einstein statistics; Pauli’s exclusion principle does not apply to bosons and bosons may condense [☛ Bose condensate]. By the spin-statistics theorem (in Lorentz-covariant models), physical particles whose mathematical representatives transform as tensor representations of the Lorentz group are bosons. The possible values of bosonic wave-functions and fields are (ordinary) commuting numbers (“c-numbers”).

BRST quantization A procedure (by Carlo M. Becchi, Alain Rouet and Raymond Stora, and separately by Igor V. Tyutin) of constructing a quantum theory from an originally classical field theory with a gauge symmetry, in which the gauge symmetry reduces to a BRST symmetry and counterterms are added to the Lagrangian density that are invariant with respect to the BRST symmetry, although not with respect to the original (classical) gauge symmetry. As a gauge symmetry is realized in quantum theory by imposing constraints (that the physical states are invariant under the action of the symmetry) and these constraints close an algebra, BRST quantization is a canonical generalization of the Dirac quantization with constraints of the first class in Dirac's classification [445, 425, 345]; see also the texts [555, 484, 496, 589, 590]. [☛ Dirac quantization; canonical quantization]

BRST symmetry A reduction of a gauge symmetry where the parameters in a gauge transformation are replaced by *ghost fields*: functions of spacetime that have the opposite statistics from the original parameters but transform identically as the original parameters under the action of both the gauge and the Lorentz transformations. For example, Yang–Mills gauge theories have ordinary (commutative) scalar functions as gauge parameters. In the corresponding BRST symmetry, to the system is added a pair of canonically conjugate *anticommutative* scalar fields that otherwise, in every other aspect, transform identically as the original parameters of the given gauge transformation. Interactions of these ghost fields with other fields are determined precisely so that they cancel the contributions of the unphysical components in the gauge fields [44]; see also the texts [268, 555, 484, 496, 589, 590]. [☛ ghost fields; nonphysical components]

Bundle [☞ vector bundle]

BV-quantization A contemporary version of canonical quantization (by Jean Zinn–Justin, then by Igor Batalin and Grigori Vilkovisky) in the Lagrangian formalism, which generalizes the evolution of the canonical–Dirac–BRST quantization to the general case when the constraints do not close the structure of an algebra [41, 345]; see also the texts [555, 484]. [☛ BRST quantization; Dirac quantization; canonical quantization]

Canonical quantization Also known as the second quantization; the adjective "canonical" stems from using the canonical Hamiltonian formalism of classical physics and its quantum reinterpretation, where the relations between observables in a given model are preserved as well as possible, and with a formal replacement of the Poisson brackets by commutators. Changes in these relations, e.g., if the Poisson bracket $\{\mathcal{A}, \mathcal{B}\} = \mathcal{C}$ upon canonical quantization becomes $[\mathsf{A}, \mathsf{B}] = \mathsf{C} + \Delta$, the additional term Δ is one of the measures of this *anomaly*. [☛ anomaly]

Cartesian product Also known as the *direct product*: for two sets X and Y, the Cartesian product is the set of all ordered pairs:

$$X \times Y := \big\{(x, y) \,:\, x \in X,\ y \in Y\big\}. \tag{B.1}$$

Cauchy sequence Given a metric space (a set of points x_i with a well-defined distance function $d(x_i, x_j)$ between any two points), this is a sequence of points $x_1, x_2, \ldots$, where

$$d(x_i, x_j) < \epsilon, \quad \forall i, j > N, \tag{B.2}$$

for some predefined integer N and positive real number (tolerance) ϵ. That is, all points sufficiently far up the sequence are closer than ϵ to each other.

Chirality The eigenvalue of the operator $\widehat{\boldsymbol{\gamma}}$. A particle is said to have a well-defined chirality if its wave-function is an eigenfunction of this operator. The operators $\frac{1}{2}[\mathbb{1} \pm \widehat{\boldsymbol{\gamma}}]$, with the $\widehat{\boldsymbol{\gamma}}$-matrix defined in Appendix A.6.1, project to spin-$\frac{1}{2}$ particles of chirality $\pm\frac{1}{2}$. By construction, chirality is Lorentz-invariant. However, as $\widehat{\boldsymbol{\gamma}}$ anticommutes with the Dirac operator $\boldsymbol{\gamma}^\mu \partial_\mu$ and commutes with the mass,

$$[i\hbar\boldsymbol{\gamma}^\mu\partial_\mu - mc\mathbb{1}]\, \tfrac{1}{2}[\mathbb{1} \pm \widehat{\boldsymbol{\gamma}}] \not\propto \tfrac{1}{2}[\mathbb{1} \pm \widehat{\boldsymbol{\gamma}}]\, [i\hbar\boldsymbol{\gamma}^\mu\partial_\mu - mc\mathbb{1}], \tag{B.3}$$

and the chirality of a massive particle is not a constant.

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

CM system For a system of particles located at the positions $\vec{r}_i$ and having the masses m_i, the position and velocity of the center of mass are, by definition,

$$\vec{r}_{CM} := \frac{\sum_i m_i \vec{r}_i}{\sum_i m_i}, \qquad \vec{v}_{CM} := \frac{\sum_i m_i \vec{v}_i}{\sum_i m_i}. \tag{B.4}$$

A coordinate system where $\vec{v}_{CM} = 0$ is called the center of momentum frame, where $\vec{r}_{CM}$ need not vanish; a coordinate system where additionally also $\vec{r}_{CM} = 0$ is called the center of mass system, or "CM-system" for short.

Codimension For a subspace $X \subset Y$, $\text{cod}(X \subset Y) := \dim(Y) - \dim(X)$. If the subspace X is defined by means of a system of algebraic equations, near every point $x \in X$, that system must have $\text{cod}(X \subset Y)$ independent equations.

Codomain For a mapping $f : X \to Y$, the collection of elements Y wherein the map points, and wherein the *values* of f and its *image* lie; $f(x) = y \in Y$ for all $x \in X$.

Cokernel For a linear mapping $f : X \to Y$ of a vector space X into Y, the *cokernel* of f consists of the equivalence classes $\text{cok}(f) := \{[y \simeq y + f(x)] : \ x \in X, \ y \in Y\}$.

Color In the context of elementary particles, the 3-dimensional $SU(3)_c$ charges of quarks, such that baryons consist of three quarks with one of the three linearly independent colors ("red," "yellow," "blue") each, so that the baryon is "colorless," or more precisely, $SU(3)_c$-invariant. Owing to the ubiquity of computer graphics, the so-called subtractive color system is ever more familiar, but we adopt the familiar additive color system. Here, red and yellow produces orange, and its mix with blue produces ***colorless***, i.e., black. The opposite (anti-)colors of primary colors are: anti-red = green, anti-yellow = purple, anti-blue = orange; the mixture of any color and its anti-color produces ***colorless***. Because of this regularity the name *color* is convenient as a mnemonic crutch for adding $SU(3)_c$ vectors [☞ Appendix A.4].

Compact space A *topological space* [☛ topological space] X where every open neighborhood (and so also the whole X) may be covered by a finite number of open neighborhoods is called quasi-compact. A topological space where every two distinct points have some non-intersecting neighborhoods,

$$\forall x \neq x' \in X, \quad \exists U, U' \subset X : \quad U \ni x, \ \ U' \ni x', \ \ U \cap U' = \varnothing \tag{B.5}$$

is called Hausdorff. A Hausdorff space that is also quasi-compact is compact. In practice in theoretical physics, it is crucial that compact spaces have a well-defined *size*, so that compact spaces may be chosen to be smaller (or larger) than a given size/length.

Compactification The procedure where a non-compact topological space X is added to a topological space Y of strictly lesser dimension, so that $X^c := (X \cup Y)$ is compact. The simplest example is $S^1 = \mathbb{R}^1 \cup \{\text{point}\}$, where a point "at infinity" was added to the open line ($\mathbb{R}^1$), so as to obtain the circle (S^1).

Concrete applications of this procedure within the present subject stem from the proposal originally made by Gunnar Nordstrøm, in 1914, whereby the spatial dimension of the form of an open and infinitely large line, $\mathbb{R}^1$, is replaced by a closed, compact and small circle, S^1. The proposal was rediscovered by Theodor F. E. Kałuża in 1919 (published in 1921) and also Oscar Klein in 1921. The latter two publications being generally known, this is typically called "Kałuża–Klein compactification." Symmetries of the compactified space result in Yang–Mills gauge symmetries in the non-compact spacetime. The special case when the compact space is a Calabi–Yau manifold is called "Calabi–Yau compactification." As Calabi–Yau spaces of more than one complex dimension do not have continuous symmetries, Calabi–Yau compactification does not give rise to any gauge symmetry, and in fact typically reduces what gauge symmetry there was prior to compactification; see Section 11.3.1.

Complex structure [☛ conjugation]
Commutative [☛ abelian]
Conjugation is a mapping of one (generalized) *complex structure* into its equivalent partner. Most generally, a ***complex structure*** is specified by an operation $\hat{\mathcal{I}}$, the two-fold repetition of which results in a sign change: $\hat{\mathcal{I}} \circ \hat{\mathcal{I}} = -\mathbb{1}$. Therefore, $-\hat{\mathcal{I}}$ is always also a complex structure, distinct from $\mathcal{I}$ but equivalent to it for all purposes, and all complex structures always occur in such equivalent pairs.

Complex conjugation Every rule by which a pair of real numbers (x, y) is assigned a complex number z has a conjugate rule. For example, *relative* to the definition $z := (x + iy)$, $z^* := (x - iy)$ is the ***complex conjugation*** of the complex number z. Operatively, complex conjugation changes $i \to -i$. The analogous situation holds also for matrices, functions, operators, etc.

Hermitian conjugation of matrices is the combination of complex conjugation (of every element) of the matrix with its transposition: $(a_{ij})^\dagger := a^*_{ji}$. [☛ Digression 10.2 on p. 360]

Dirac conjugation of a Dirac spinor Ψ is the Hermitian conjugation combined with right-multiplication by the $\boldsymbol{\gamma}^0$ matrix: $\overline{\Psi} := \Psi^\dagger \boldsymbol{\gamma}^0$. Correspondingly, the Dirac conjugate of the operator R is $\overline{R} := (\boldsymbol{\gamma}^0)^{-1} R \boldsymbol{\gamma}^0$. For a Cartesian basis of $\boldsymbol{\gamma}$-matrices with the metric tensor (3.19), it follows that $(\boldsymbol{\gamma}^0)^{-1} = \boldsymbol{\gamma}^0$ so $\overline{R} = \boldsymbol{\gamma}^0 R \boldsymbol{\gamma}^0$, which agrees with the definition (5.132).

Contact interaction Interaction that requires that all participants in the interaction are localized in the same spacetime point – akin to the collision of two marbles. All elementary processes in the Standard Model are contact interactions. For example, the emission and the absorption of a (virtual) photon by an electron requires that the "incoming" electron in a spacetime point turn into the "outgoing" electron and that the photon in this interaction is emitted from or absorbed at that same point. The Yukawa interaction is analogous, except that a scalar particle is emitted or absorbed instead of a photon. The Fermi interaction is also analogous, except that here two fermions collide in a spacetime point from which then two other fermions emerge, or one fermion decays into three fermions, all emitted from the same spacetime point.

Contravariant vector A vector the components of which, $A^\mu(\mathrm{x})$, are transformed as

$$A^\mu(\mathrm{x}) = \left(\frac{\partial x^\mu}{\partial y^\nu}\right) A^\nu(\mathrm{y}) \tag{3.11c}$$

by the coordinate system transformation $\mathrm{x} \to \mathrm{y}$.

Coset [☛ Appendix A.1.1.]

Cotangent bundle The vector bundle $\mathscr{T}^*_{\mathscr{X}} := E(\mathscr{X}; T^*_x(\mathscr{X}); \pi)$ where $T^*_x(\mathscr{X})$ is the cotangent space of the space $\mathscr{X}$ at the point $x \in \mathscr{X}$. If x^μ are local coordinates in the space $\mathscr{X}$ at the given point, then $T^*_x(\mathscr{X})$ may be represented as the formal vector space of linear combinations $\omega_\mu \mathrm{d}x^\mu$.

Coulomb field, potential A stationary electric charge is surrounded by the constant Coulomb electrostatic field, $\vec{E}$; $q_0\vec{E}$ is the force that acts upon the probing particle of charge q_0. For the same situation, $\vec{E} = -\vec{\nabla}\Phi$, where Φ is the Coulomb potential; $q_0\Phi$ is the potential energy of the probing charge q_0 in the field $\vec{E}$. It follows from Gauss's law that the Coulomb field of a point-like charge is $\vec{E} \propto 1/r^{d-1}$, where d is the dimension of the space and r the distance between the source of the field and the place where the field is measured; also, $\Phi \propto 1/r^{d-2}$.

Covariant derivative A measure of the amount of change in the "overall value" of a generalized function $\boldsymbol{F}$ owing to a change of one of the arguments of $\boldsymbol{F}$ in the limiting case when the change in the argument is infinitesimal and tends to zero. For a real scalar (invariant) function, the "overall value" is simply the "value" or intensity, and the covariant derivative of such a function is the same as the partial derivative. However, for more general functions $\boldsymbol{F}$ that take values in a multi-dimensional space, such as spacetime itself or some abstract space, the covariant derivative also takes into account that the space of values of $\boldsymbol{F}$ may well change over the space of arguments. This then additionally changes the "overall value" (both the intensity and the "direction") of $\boldsymbol{F}$ at an infinitesimally close neighboring value of the argument. The covariant derivatives therefore have the general form $D := \partial + \boldsymbol{\Gamma}$, where $\boldsymbol{\Gamma}$ is the gauge potential and encodes the variation in the space of values of $\boldsymbol{F}$. [☛ gauge potential, gauge field]

Covariant vector A vector the components of which, $B_\mu(\mathrm{x})$, are transformed as

$$B_\mu(\mathrm{x}) = \Big(\frac{\partial y^\nu}{\partial x^\mu}\Big)\, B_\nu(\mathrm{y}) \tag{3.11d}$$

by the coordinate system transformation $\mathrm{x} \to \mathrm{y}$.

Covering For a given (topologial) space X, the n-fold (finite) cover Y is a space for which there exists an n–1 mapping $\pi : Y \to X$ such that for every point $x \in U \subset X$, where U is any open neighborhood in X, there exist exactly n points and non-intersecting open neighborhoods $y_i \in V_i \subset Y$, such that $\pi(y_i) = x$ and $\pi(V_i) = U$. That is, π is a continuous surjection. The points y_i are called the π-inverse images of x, i.e., $\pi^{-1}(x) = \{y_1, y_2, \dots\}$.

Curvature Given a space $\mathscr{X}$ over which the functions $f(x^\mu)$ are defined locally, i.e., in sufficiently small open neighborhoods, $f(x^\mu)$ is unambiguously defined. Let D_μ be local derivatives that (in sufficiently small open neighborhoods) correctly compute the difference $\mathrm{d}x^\mu(D_\mu f) = f(x^\mu{+}\mathrm{d}x^\mu) - f(x^\mu)$. Then, in general, the relations

$$\big[\, D_\mu \,,\, D_\nu \,\big] = \boldsymbol{T}_{\mu\nu}{}^\rho D_\rho + \boldsymbol{R}_{\mu\nu} \tag{B.6}$$

define the ***torsion*** $\boldsymbol{T}_{\mu\nu}{}^\rho$ and the ***curvature*** $\boldsymbol{R}_{\mu\nu}$ of the space $\mathscr{X}$. These two (local) structures specify the (local) geometry of the space $\mathscr{X}$ and a class of functions $f(x^\mu)$ over this space.

In examples where $\mathscr{X}$ is spacetime and $f(x^\mu)$ a complex wave-function representing a lepton or a quark, the torsion vanishes, and $\boldsymbol{R}_{\mu\nu}$ is the Yang–Mills gauge field (denoted $\mathbb{F}_{\mu\nu}$ [☞ Chapters 5 and 6]). The torsion vanishes also when $f(x^\mu)$ represents a tensor over spacetime $\mathscr{X}$, in which case is $\boldsymbol{R}_{\mu\nu}$ the Riemann tensor [☞ Chapter 9]. In turn, when $D_\alpha, \overline{D}_{\dot\alpha}$ are super-derivatives (10.68), so the commutator in the relation (B.6) is replaced by an anticommutator, the curvature vanishes and the torsion does not [☞ relation (10.69), which holds for the extended basis of super-spacetime derivatives $\{D_\alpha, \overline{D}_{\dot\alpha}, \partial_\mu\}$]. Finally, in the theory of Lie groups, the Lie group itself is a differentiable space where the derivatives are closely related to the generators Q, and their commutator, akin to (A.70), defines the *structure constants* of the Lie group as the torsion and where the curvature vanishes.

Dirac quantization The development of general canonical quantization for systems in which there exist constraints, and the specification how to treat these constraints in quantum theory so they remain satisfied throughout the evolution of the system in time; see Digression 11.7 on p. 420, the texts [64, 445, 425], as well as Dirac's book [134]. Dirac's procedure proves the fundamental equivalence between Heisenberg's "matrix mechanics" and Schrödinger's "wave mechanics" and connects the ideas from both approaches. [☛ canonical quantization]

Direct product [☛ Cartesian product]

Domain For a map $f : X \to Y$, this is X, the collection of elements that are being mapped by f; $X := \{x : f(x)$ is well defined$\}$.

Einstein–Rosen bridge A wormhole that connects the inside of the event horizon of one of two Schwarzschild black holes with the inside of the event horizon of another black hole of the same type. [☛ wormhole]

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

Energy–momentum (4-momentum) transfer In collisions $A + B \to A' + \cdots$, where B is initially a target at rest and A and A' the incoming and outgoing probe,[1] $\mathrm{q} := (\mathrm{p}_A - \mathrm{p}_{A'})$ is the 4-momentum that the probe transfers to the target. In elastic collisions, $A + B \to A' + B'$, we have that $\mathrm{q} = (\mathrm{p}_{B'} - \mathrm{p}_B)$.

Equivalence A binary relation $\sim$ between elements of a set A is an equivalence if and only if it is (**1**) reflexive ($a \sim a$), (**2**) symmetric (if $a \sim b$ then $b \sim a$), (**3**) transitive (if $a \sim b$ and $b \sim c$ then $a \sim c$). An equivalence class is a subset of A consisting of elements that are all equivalent to each other; different equivalence classes are disjoint subsets of A, and their union equals A.

Euler characteristic Denoted $\chi_E(\mathscr{X})$, the Euler (or the Euler–Poincaré) characteristic is the topological invariant of the topological space $\mathscr{X}$. If $\mathscr{X}$ is a real 2-dimensional surface that has a triangulation (an approximation by a network of finitely many triangles), $\chi_E(\mathscr{X}) = k_0 - k_1 + k_2$, where k_0 is the number of vertices (corners), k_1 the number of edges and k_2 the number of triangles. A generalization exists also to higher-dimensional spaces (using a generalization of triangles): $\chi_E(\mathscr{X}) = \sum_{i=0}^{\dim X}(-1)^i k_i$, where k_0, k_1, k_2 are defined as for surfaces, k_3 the number of (exclusively tetrahedral) 3-dimensional elements, etc.

Extremal black hole A nontrivial solution of the Einstein equations, such as the Reissner–Nordstrøm solutions (9.61) where the two horizons coincide, $2r_q = r_S$, and which is marginal between the solutions where the singularity is screened by the event horizon and the solutions where it is not, i.e., solutions with a naked singularity.

Fermion By definition, a particle (as well as its mathematical representatives: wave-functions, creation and annihilation operators, or fields) that obeys Pauli's exclusion principle (two fermions cannot be in the same quantum state) and therefore also the Fermi–Dirac statistics. Owing to the spin-statistics theorem (in Lorentz-covariant models), physical particles whose mathematical representatives transform as spinorial representations of the Lorentz group are fermions. Fermionic wave-functions and fields have values that are anticommuting "numbers" ("a-numbers").

Fibration The space obtained by generalizing the tensor product of two spaces, where one of the factors in the product changes "along" the other factor. The type of that change (continuous, smooth, analytic, complex-analytic, …) distinguishes the various fibrations. Even the topology, i.e., homotopy of the variable factor may change, i.e., this factor may change discontinuously. [☛ homotopy class, Figure 11.7 on p. 427]

Field (mathematics) A collection of elements, $\Bbbk$, for which two operations, # and $*$, are defined so that:

1. $(\Bbbk, \#)$ is an abelian (commutative) group, with $e \in \Bbbk$ the neutral element;
2. $(\Bbbk \smallsetminus \{e\}, *)$ is an abelian (commutative) group;
3. the distribution rules $a * (b \# c) = (a * b) \# (a * c)$ and $(a \# b) * c = (a * c) \# (b * c)$ hold.

Field (physics) A function over spacetime. A scalar field is a function the values of which are scalars, a vector field is a function the values of which are vectors, etc. By a "*gauge field*," however, one means the concrete fields such as the electric and magnetic fields, and their generalizations to other gauge models. [☛ gauge field] Variations/perturbations in a field are quantized in quantum physics. [☛ quantum]

Flavor The type of quark – distinguished by their masses and various charges, see the tabulation (2.44a). These are eigenstate of the free (propagation) Hamiltonian, and flavor ranges over *up*, *down*, *strange*, *charm*, *beauty* and *top*.

Gauge fields In the most familiar example, electromagnetism, these are the electric and the magnetic fields, which jointly form Maxwell's tensor $F_{\mu\nu}$ [☞ relations (5.73)]. More generally,

[1] A and A' are one and the same particle, with changed kinematical parameters: energy, linear momentum and angular momentum, including spin.

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

Yang–Mills gauge fields are the components of the matrix-valued tensor $\mathbb{F}_{\mu\nu}$ [☞ definition (6.15)], and for gravity these are the components of the Riemann tensor (9.30). In the most general case, gauge fields are defined, up to multiplicative constants, as the result of computing $[D_\mu, D_\nu]$, where D_μ are the correspondingly gauge-covariant derivatives, so $[D_\mu, D_\nu]$ is a measure of the non-commutativity of the changes of the considered generalized (complex-, vector-, tensor-, spinor-, matrix-, Lie-algebra-, … valued) functions, i.e., the curvature of the space of such generalized functions. [☛ covariant derivative]

Gauge potential In the most familiar example, electromagnetism, these are the scalar and the vector potentials that jointly form the 4-vector A_μ [☞ relations (5.73)] and represent the difference between the covariant and the partial derivative [☞ definition (5.13)]. More generally, Yang–Mills gauge potentials form a matrix-valued 4-vector $\mathbb{A}_\mu$ [☞ definition (6.6a)], and for gravity these are the Christoffel symbols (9.17). In the most general case, the gauge potential is the difference between the gauge-covariant and the partial derivative: $\boldsymbol{\Gamma} = D - \partial$. [☛ covariant derivative, potential]

Geodesic completeness The property of a given coordinate system with the given metric tensor that the limiting points of all geodesic lines (9.48) are within the range of those coordinates. A typical nontrivial example is the surface of a torus, for which we choose the coordinates (x, y), where x parametrizes the "little circle" so $x \simeq x + 2\pi R_1$, and y parametrizes the "big circle" so $y \simeq y + 2\pi R_2$, with $R_2 \geqslant R_1$. The coordinate system (x, y) is thus geodesically complete. As a counter-example, consider the "northern" stereographic projection of a sphere to the (x, y)-plane, so that the south pole corresponds to the coordinate origin and the equator to the circle of unit radius centered at the coordinate origin. Then geodesic lines on the sphere that contain the north pole correspond to geodesic lines in the plane that contain the point at infinity – which is not within the range of the coordinates. Such geodesic lines are thus incomplete or even disconnected, so that the coordinate system (x, y) with any Euclidean metric is geodesically incomplete as a description of a sphere.

Geometric quantization The process of constructing a quantum theory from the original classical theory, which uses the symplectic structure ω of the phase space Φ of the classical theory [288, 173, 579, 56]. Observables in classical theory are simply real functions $\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$ over Φ. Geometric quantization is based on the introduction of a ω-compatible *polarization* $\pi(\Phi)$. In physics practice, π denotes the concrete choice of the half of the coordinates in the phase space Φ, which are the canonical coordinates, q^i, for which the ω-complementary half of the coordinates over Φ play the role of canonically conjugate momenta, p_i. With that standard notation, the symplectic structure is simply given by the Poisson brackets $\omega(\mathcal{A}, \mathcal{B}) := \frac{\partial \mathcal{A}}{\partial q^i}\frac{\partial \mathcal{B}}{\partial p_i} - \frac{\partial \mathcal{A}}{\partial p_i}\frac{\partial \mathcal{B}}{\partial q^i}$. That same polarization produces the quantum observables $A = \pi(\mathcal{A})$, $B = \pi(\mathcal{B})$, etc. The difference

$$\Delta := [\pi(\mathcal{A}), \pi(\mathcal{B})] - \pi\big(\omega(\mathcal{A}, \mathcal{B})\big) \tag{B.7}$$

is one of the measures of *anomaly*. [☛ anomaly]

Geometrization of physics The process by which physics is increasingly described in terms of geometry. At its simplest, this is the dual interpretation of the geodesic equation either as a bending of trajectories owing to spacetime curvature (9.48) or owing to the action of a gravitational force (9.49). At a rather more comprehensive level, in string theory models compactified on a space $\mathscr{Y}$, many of the physical properties of the effective particle physics model are derived as geometrical and topological characteristics of $\mathscr{Y}$; see discussion on p. 402 and in Section 11.3.1.

Ghost field Of the four components of the gauge 4-vector potential $\mathbb{A}_\mu$, only two correspond to degrees of freedom with a physical meaning. It turns out that it is possible to introduce two (anticommuting scalar) "ghost fields," the detailed kinematics and dynamics of which are

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

chosen precisely so as to cancel the extraneous contributions of the two unphysical degrees of freedom in the 4-vector $\mathbb{A}_\mu$ [441, 425, 555, 484, 496, 589, 590]. The gauge symmetry is thereby reduced to the nilpotent BRST symmetry.

Gluon The particle (quantum) that mediates the strong interaction. Gluons interact with each other as well as with quarks and antiquarks, which they bind into hadrons. The interaction between hadrons is then a residual interaction, just as the molecular forces between electrically neutral atoms are modeled as dipole–dipole and higher order electromagnetic interactions [☞ Section 6.1.1].

Gödel's incompleteness theorem This theorem proves that no axiomatic system that is sufficiently complex to contain arithmetics can be both complete and self-consistent. Gödel's proof is constructive, and shows that within all such self-consistent axiomatic systems it is explicitly possible to construct a statement that can neither be proven nor disproven within the given axiomatic system. Therefore, either that statement or its logical negation may *always* be added to the axiomatic system as a new axiom, and this extensibility never stops [211, 376]. Although Gödel constructed a particular undecidable statement in his proof, and expressly for the purpose of proving the theorem, it does follow that there exist infinitely many such undecidable statements – and some of those, within physics as a formal axiomatic system, are bound to be of interest. [☛ Appendix B.3]

Gram–Schmidt procedure In a vector space V, equipped with a finite scalar product, i.e., where $\langle a|b\rangle < \infty$ for every $a, b \in V$, the Gram–Schmidt procedure produces an *orthonormal basis*:

1. Pick an element $a \in V$ and define $\alpha_1 := a/\sqrt{\langle a|a\rangle}$ and set $k = 1$.
2. If there is some $b \in V$ that is linearly independent from $\alpha_i \in \mathscr{B}_V := \{\alpha_1, \dots, \alpha_k\}$,
   (a) Define $\alpha_{k+1} := \sum_{i=1}^{k} c_i\alpha_i + c_{k+1}b$.
   (b) Determine $\{c_1, \dots, c_{k+1}\}$ so that
      i. $\langle\alpha_{k+1}|\alpha_i\rangle = 0$, for all $i = 1, \dots, k$,
      ii. and $\langle\alpha_{k+1}|\alpha_{k+1}\rangle = 1$.
   (c) Increase k by one ($k \mapsto k+1$), and return to step 2.
3. The basis for the vector space V is $\mathscr{B}_V = \{\alpha_1, \dots, \alpha_k\}$ and $\dim(V) = k$.

Group A collection of elements G equipped with a binary operation $\star$ that satisfies the four axioms [☛ Appendix A.1.1]:

closure $\forall a, b \in G$, $(a \star b) \in G$;
associativity $\forall a, b, c \in G$, $(a \star b) \star c = a \star (b \star c)$;
neutral element $\exists e \in G$ such that $\forall a \in G$, $a \star e = a = e \star a$;
inverse element $\forall a \in G$, $\exists a^{-1} \in G$ such that $a^{-1} \star a = e = a \star a^{-1}$.

That is, a group is an invertible monoid.

Groupoid [☛ magma]

Hadron A particle that interacts by means of the strong nuclear force (at $\sim 10^{-23}$ s) and can be detected as an isolated particle; e.g., a proton or a pion.

Hausdorff space A topological space in which distinct points have disjoint neighborhoods. Most variables typically considered in physics models span/form Hausdorff spaces. Examples of non-Hausdorff spaces include bifurcating (Y-shaped) 1-dimensional lines such as the Feynman diagrams (3.130)–(3.131) and the left-hand side of Figures 11.3 on p. 411 and 11.4 on p. 412. [☛ topological space]

Helicity The eigenvalue of the operator $\hat{p}\cdot\vec{S}/\hbar$, i.e., the projection of spin in the direction of motion of the particle, in units of $\hbar$. As massless particles move at the speed of light in vacuum, their helicity is Lorentz-invariant and equals their chirality.

Hermitian conjugation [☛ Digression 10.2 on p. 360]

Homotopy class Geometric objects that can be continuously transformed one into another form a homotopy class of such objects; different objects in the same homotopy class are homotopy

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

equivalents of each other. Continuous interpolation between two homotopy equivalent objects is called the homotopy (between those two objects). Thus is the surface of a sphere a homotopy equivalent of the surface of a cube and a tetrahedron for example, but not of a torus or a pretzel.

Hypersurface The subspace $X \subset Y$ is a hypersurface if the codimension $\mathrm{cod}(X \subset Y) = 1$; near every point $x \in X$, the subspace $X \subset Y$ is specified by a single constraint.

Image For a mapping $f : X \to Y$, the f-image of the space X is the collection of points in Y obtained by mapping the points of X: $\mathrm{im}(f) = f(X) = \{f(x) = y \in Y :\ x \in X\}$.

Injection A "1–1" (one-to-one) mapping $f : X \hookrightarrow Y$, such that for every $a \in A$ there is precisely one $y = f(x) \in Y$.

Isometry A symmetry of a space $\mathscr{X}$ that leaves the metric on $\mathscr{X}$ unchanged.

Isomorphism Bijective homomorphism, i.e., a bijection $f : X \to Y$ for which both f and f^{-1} preserve the algebraic structure of the objects X and Y, and so are ***homomorphisms***. For example, if X and Y are groups, the f-image of every group axiom in X results in the corresponding group axiom in Y, and vice versa. We write $X \cong Y$.

KamiokaNDE The Kamioka Nucleon Decay Experiment, run at the Kamioka Observatory, Institute for Cosmic Ray Research, near the Kamioka section of the city of Hida, Japan. KamiokaNDE was initially designed to detect proton decay, but was successfully used to detect solar and atmospheric neutrinos, through upgrades known as KamiokaNDE-II, Super-KamiokaNDE, Super-KamiokaNDE-II and -III.

Kernel Elements of a vector space X that a linear mapping $f : X \to Y$ maps to $0 \in Y$ form the ***kernel*** of the linear mapping f, denoted $\ker(f) := \{x \in X,\ f(y) = 0 \in Y\}$. In other words, $\ker(f)$ consists of the elements of the vector space X annihilated by the mapping f.

Kronecker product The special case of the *tensor product* for matrices of arbitrary size, so including also column-matrices and row-matrices. The result of the Kronecker product is the block-matrix:

$$\mathbb{A} = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}, \quad \mathbb{B} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \quad \text{then} \quad \mathbb{A} \otimes \mathbb{B} = \begin{bmatrix} a\begin{bmatrix} \alpha \\ \beta \end{bmatrix} & b\begin{bmatrix} \alpha \\ \beta \end{bmatrix} & c\begin{bmatrix} \alpha \\ \beta \end{bmatrix} \\ d\begin{bmatrix} \alpha \\ \beta \end{bmatrix} & e\begin{bmatrix} \alpha \\ \beta \end{bmatrix} & f\begin{bmatrix} \alpha \\ \beta \end{bmatrix} \end{bmatrix} = \begin{bmatrix} a\alpha & b\alpha & c\alpha \\ a\beta & b\beta & c\beta \\ d\alpha & e\alpha & f\alpha \\ d\beta & e\beta & f\beta \end{bmatrix}. \tag{B.8}$$

Note that $\mathbb{B} \otimes \mathbb{A} \neq \mathbb{A} \otimes \mathbb{B}$:

$$\mathbb{B} \otimes \mathbb{A} = \begin{bmatrix} \alpha\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \\ \beta\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \alpha a & \alpha b & \alpha c \\ \alpha d & \alpha e & \alpha f \\ \beta a & \beta b & \beta c \\ \beta d & \beta e & \beta f \end{bmatrix} = \begin{bmatrix} a\alpha & b\alpha & c\alpha \\ d\alpha & e\alpha & f\alpha \\ a\beta & b\beta & c\beta \\ d\beta & e\beta & f\beta \end{bmatrix} \neq \begin{bmatrix} a\alpha & b\alpha & c\alpha \\ a\beta & b\beta & c\beta \\ d\alpha & e\alpha & f\alpha \\ d\beta & e\beta & f\beta \end{bmatrix} = \mathbb{A} \otimes \mathbb{B}. \tag{B.9}$$

Kronecker symbol The index representation of the identity matrix

$$\delta^i_j := \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases} \tag{B.10}$$

allows the generalizations after the pattern:

$$\delta_{[k\ell]}^{\,ij} := \tfrac{1}{2}\big(\delta^i_k\delta^j_\ell - \delta^i_\ell\delta^j_k\big), \qquad \delta_{(k\ell)}^{\,ij} := \tfrac{1}{2}\big(\delta^i_k\delta^j_\ell + \delta^i_\ell\delta^j_k\big), \tag{B.11a}$$

$$\delta_{[\ell mn]}^{\,ijk} := \tfrac{1}{3!}\big(\delta^i_\ell\delta^j_m\delta^k_n - \delta^i_\ell\delta^j_n\delta^k_m + \delta^i_n\delta^j_\ell\delta^k_m - \delta^i_n\delta^j_m\delta^k_\ell + \delta^i_m\delta^j_n\delta^k_\ell - \delta^i_m\delta^j_\ell\delta^k_n\big), \tag{B.11b}$$

$$\delta_{(\ell mn)}^{\,ijk} := \tfrac{1}{3!}\big(\delta^i_\ell\delta^j_m\delta^k_n + \delta^i_\ell\delta^j_n\delta^k_m + \delta^i_n\delta^j_\ell\delta^k_m + \delta^i_n\delta^j_m\delta^k_\ell + \delta^i_m\delta^j_n\delta^k_\ell + \delta^i_m\delta^j_\ell\delta^k_n\big), \quad \text{etc.,} \tag{B.11c}$$

which are also called (anti-)symmetrized Kronecker symbols.

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

Lepton A particle that does not interact by means of the strong nuclear force (at $\sim 10^{-23}$ s); e.g., the electron.

Levi-Civita symbol The index representation of the permutation symbol

$$\varepsilon_{i_1 \cdots i_n} := \begin{cases} +1, & \text{if the order } i_1, \dots, i_n \text{ is an even permutation of } 1, 2, \dots, n, \\ -1, & \text{if the order } i_1, \dots, i_n \text{ is an odd permutation of } 1, 2, \dots, n, \\ 0, & \text{otherwise.} \end{cases} \tag{B.12}$$

We also define $\varepsilon^{i_1 \cdots i_n} := \varepsilon_{i_1 \cdots i_n}$. (Some Authors prefer using a definition such that $\varepsilon^{i_1 \cdots i_n} := -\varepsilon_{i_1 \cdots i_n}$, for numerical convenience in some computations.) The key relation between the Levi-Civita and the Kronecker symbols is

$$\begin{aligned} \varepsilon^{i_1 \cdots i_n} \varepsilon_{j_1 \cdots j_n} &= \delta^{\,i_1 \cdots i_n}_{[j_1 \cdots j_n]}, \\ &= \tfrac{1}{n!} \Big(\delta^{i_1}_{j_1} \cdots \delta^{i_{n-1}}_{j_{n-1}} \delta^{i_n}_{j_n} - \delta^{i_1}_{j_1} \cdots \delta^{i_{n-1}}_{j_n} \delta^{i_n}_{j_{n-1}} + \cdots \quad (n! \text{ permutations, total}) \Big). \end{aligned} \tag{B.13}$$

Lie group [☛ Appendix A.1.1]

Luxon a particle that travels through vacuum at the speed of light in vacuum, c, and has no mass. All mediators of gauge interactions that correspond to unbroken gauge symmetries are luxons.

Magma (groupoid) A collection of elements M equipped with a *closed* binary operation $\star$, i.e., $\forall a, b \in M$, $(a \star b) \in M$.

Manifold A space where every sufficiently small neighborhood of every point is isomorphic to the flat space $\mathbb{R}^n$, where n is the dimension of the manifold. A manifold is everywhere smooth and the tangent space at every point is a copy of $\mathbb{R}^n$.

Mass shell In the 4-dimensional space of 4-momentum, the "mass shell" for a particle of mass m is the subspace defined by the relation $E^2 - \vec{p}^2 c^2 = m^2 c^4$. For $m^2 > 0$ (ordinary particles and antiparticles), this is the two-component hyperboloid, where $E = \pm\sqrt{m^2 c^4 + \vec{p}^2 c^3}$ on both "shells." For $m = 0$ (photons, gluons and gravitons), this is the "light cone" the two portions of which touch in the point $(E/c, \vec{p}) = (0, \vec{0})$. For $m^2 < 0$ (tachyons), this is the single-component hyperboloid.

Meson Since the acceptance of the quark model in 1973, a bound state of a quark and an anti-quark. Originally, a particle that interacts by means of the strong nuclear force (at $\sim 10^{-23}$ s), can be detected as an isolated particle, and has a mass that is between the electron mass and the proton mass; e.g., $\pi^\pm, \pi^0$.

Minimal coupling The coupling between matter and interaction field that occurs by the interaction field modifying the spacetime derivative of the matter field. The gauge principle introduces only minimal coupling [☛ Chapters 5–7 and 9].

For example, let $\Psi(\mathrm{x})$ represent the matter field and $A_\mu(\mathrm{x})$ the gauge potential of the interaction field. They are minimally coupled through replacing $\partial_\mu \Psi \to (\partial_\mu + i g A_\mu)\Psi$, where g is a suitable (coupling) parameter; $g = \frac{q_\Psi}{\hbar c}$ in electromagnetism, where q_Ψ is the electric charge of the matter particle represented by $\Psi(\mathrm{x})$.

Monoid A collection of elements M equipped with a binary operation $\star$ that satisfies the three axioms [☛ Appendix A.1.1]:

closure $\forall a, b \in M$, $(a \star b) \in M$;

associativity $\forall a, b, c \in M$, $(a \star b) \star c = a \star (b \star c)$;

neutral element $\exists e \in M$ such that $\forall a \in M$, $a \star e = a = e \star a$.

That is, a monoid is a semigroup with a neutral element.

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

Multipole expansion The expansion of a function over 3-dimensional flat space, in which we use spherical coordinates, over the complete system of spherical harmonics [☞ relations (4.2)–(4.4)]:

$$F(r,\theta,\phi) = \sum_{\ell,m} f_\ell^m(r)\, Y_\ell^m(\theta,\phi), \tag{B.14}$$

where

$$f_\ell^m(r) := \int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \sin\theta\, \mathrm{d}\theta\, \big(Y_\ell^m(\theta,\phi)\big)^* F(r,\theta,\phi), \tag{B.15}$$

$$\vec{\nabla}^2 F(r,\theta,\phi) = \frac{1}{r}\Big[\frac{\mathrm{d}^2}{\mathrm{d}r^2} r\, F(r,\theta,\phi)\Big] - \frac{1}{r^2}\Big[\vec{L}^2 F(r,\theta,\phi)\Big], \tag{B.16}$$

$$\vec{L}^2 Y_\ell^m(\theta,\phi) = \ell(\ell{+}1)\, Y_\ell^m(\theta,\phi); \qquad \vec{L}^2 := -\vec{\nabla}^2\big|_{r=1}, \;\; \ell \geqslant 0. \tag{B.17}$$

Notice that the coordinates θ,ϕ parametrize a 2-sphere, $S^2 = \mathbb{R}^3\big|_{r=1}$. More generally, for every compact Riemann space $\mathscr{K}$, the Laplacian $\vec{\nabla}_K^2|_{r=1}$ has a non-positive spectrum (collection of eigenvalues), and corresponding eigenfunctions, which generalize the spherical harmonics.

Noether theorem To every continuous symmetry of a physical system in classical physics, there corresponds an additive current density that satisfies the continuity equation, and produces an additive conserved charge. In quantum theory, the conserved charges are eigenvalues of generators of the corresponding symmetries, and these in turn are the momenta canonically conjugate to the canonical variables the (eigen)values of which the symmetries change. For example, the linear momentum $\vec{p}$ is the eigenvalue of the operator of linear momentum $\vec{\boldsymbol{p}}$, and also the conserved "charge" of the corresponding translations in position $\vec{r}$, which is generated by $\vec{\boldsymbol{p}} = \frac{\hbar}{i}\vec{\nabla}$ and implemented by the unitary operator $\exp\{i\vec{a}\cdot\vec{\boldsymbol{p}}/\hbar\} = \exp\{\vec{a}\cdot\vec{\nabla}\}$.

Conserved "charges" of finite symmetries are multiplicative: a product of two parity eigenfunctions is also a parity eigenfunction, with the eigenvalue that is a product of eigenvalues of the factors. Although Noether's original theorem does not apply to finite symmetries, the generalization is easy to derive. However, the operators that implement discrete symmetries may be both linear (and so unitary), and anti-linear (and then anti-unitary), such as the operator of charge conjugation: $\boldsymbol{C}(\alpha\boldsymbol{A}) = \alpha^*\boldsymbol{C}(\boldsymbol{A})$, for every operator $\boldsymbol{A}$ and constant $\alpha \in \mathbb{C}$.

Non-abelian [☛ abelian]

Non-commutative [☛ abelian]

Nonphysical components Within every Lorentz-covariant formalism, one uses only fields and operators that form complete representations of the Lorentz group. Thus, for example, gauge potentials in (3+1)-dimensional spacetime are always presented by 4-vectors, $A_\mu(\mathrm{x})$. However, only two components of this 4-vector are physically measurable, while two are not: for example, for a freely propagating field in empty space, the temporal and the longitudinal components are nonphysical. There exists no Lorentz-covariant method of isolating them from the 4-tuple (A_0, A_1, A_2, A_3). For example, the Lorenz gauge, $\eta^{\mu\nu}\partial_\mu A_\nu = 0$ specifies one *differential* relation between the 4-vector components $A_0(\mathrm{x}),\ldots,A_3(\mathrm{x})$ in a Lorentz-invariant way, which formally permits expressing one of the four components in terms of an *integral* of the derivatives of the other three components. This effectively removes one degree of freedom, but this relation is not local. However, for the removal of the other nonphysical component, there does not even exist a Lorentz-invariant gauge condition – neither algebraic nor differential. [☛ BRST quantization]

Normal subgroup A subgroup $N \subset G$ is normal if

$$\forall n \in N \subset G, \;\; \forall g \in G, \;\; gng^{-1} \in N. \tag{B.18}$$

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

Ockham's principle Also known as *Ockham's razor*, as well as the *principle of parsimony*, of *economy* and of *succinctness*, whereby from among two competing possible explanations one must choose the simpler. Although this principle is useful in research practice, one must recognize that its application depends strongly on the cultural "background": ideas and elements that are well known within one culture (and are therefore regarded simpler) may well be alien in another culture. Thus, there is a danger that the application of this principle is simply a façade of a prejudice.

Pauli's principle Two identical fermions cannot simultaneously be in the same quantum state, i.e., they cannot simultaneously occupy the same "place" in the Hilbert space.

Photon The particle (quantum) that mediates the electromagnetic interaction. Photons interact directly with quarks, antiquarks, (electrically) charged leptons (e^-, μ^- and τ^-) and also with the charged weak gauge bosons $W^\pm$ [☞ Sections 2.3.4 and 5.2.2].

Physical components In practice, physical quantities are not infrequently represented by multi-component mathematical objects such as vectors, tensors and spinors. Components that in some way may be measured experimentally (such as the transversal polarizations of the electromagnetic radiation, for example) are *physical*. [☛ nonphysical components]

Point-like The property of showing sign of neither internal structure nor spatial extension.

Potential Short for "gauge potential," this term is used as a generalization of the electro-static potential, where we have that if $\Phi(\vec{r},t)$ is the potential, then:

1. $g\,\Phi(\vec{r},t)$ is the *potential energy* of a particle with charge g when placed in the potential $\Phi(\vec{r},t)$ that interacts with this charge,
2. $-\vec{\nabla}\Phi(\vec{r},t)$ is the *(gauge) field* corresponding to the potential,
3. $-g\,\vec{\nabla}\Phi(\vec{r},t)$ is the force that the potential $\Phi(\vec{r},t)$ exerts on a particle of charge g.

In the relativistic generalization, one speaks of the "4-vector potential," $(\Phi, -c\vec{A})$, for which the *fields* are the components of the $F_{\mu\nu} := (\partial_\mu A_\nu - \partial_\nu A_\mu)$ tensor [☞ definitions (5.73)]; in the non-abelian (non-commutative) generalization the *fields* are defined as the components of the $\mathbb{F}_{\mu\nu} := [D_\mu, D_\nu]$ tensor, where $D_\mu := \partial_\mu + \frac{iq}{\hbar c}\mathbb{A}_\mu$ [☞ definition (6.15)]. Finally, in the general theory of relativity, Christoffel symbols and the connection 4-vector play the role of the *potential* and the components of the Riemann tensor are the *fields* [☞ Sections 9.2.1 and 9.2.2]. [☛ gauge potential]

Quantum In quantum physics, all material entities (matter as well as interactions thereof) are subject to quantization of the Hamilton action, which cannot vary continuously, but as integral multiples of the Planck constant, $\hbar$. Note that the "background" (settled, static, infinitely spread-out, classical, i.e., non-quantum) fields, such as the Coulomb field of a static charge distribution, are but a convenient idealization, representable by averaging over an infinite number of quanta. [☛ field (physics)]

Quotient space [☛ Appendix A.1.1]

Range (of a mapping) For a mapping $f : X \to Y$, this can variously denote either the *codomain* or the *image* of f; this ambiguity and this term are avoided herein.

Rank (of a mapping) For a mapping $f : X \to Y$, $\text{rank}(f) = \dim\big(\text{im}(f)\big) = \dim\big(f(X)\big)$.

Rank (of a tensor density) [☛ definition on p. 511]

Ring A collection of elements, $\Bbbk$, for which two operations, $\#$ and $*$, are defined so that:

1. $(\Bbbk, \#)$ is an abelian (commutative) group, with $e \in \Bbbk$ the neutral element;
2. $(\Bbbk, *)$ is a monoid (like a group, but without invertibility);
3. the distribution rules: $a*(b\,\#\,c) = (a*b)\,\#\,(a*c)$ and $(a\,\#\,b)*c = (a*c)\,\#\,(b*c)$ hold.

Semidirect product Some groups have the structure $G = H \ltimes N$, where $H \subset G$ is a subgroup, and $N \subset G$ is a normal subgroup [☛ normal subgroup]. This implies that the only common element is $N \cap H = \mathbb{1} \in G$, and that every group element $g \in G$ can be factorized as $g = h \circ n = n' \circ h'$, where $n, n' \in N$ and $h, h' \in H$. The group G is said to be an N-extension of the

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

group H; it is also true that H is isomorphic to the quotient group G/N [☞ definition (A.6) for the quotient space, which here inherits the group structure].

A well-known example is the Poincaré group, $Po(1,3) = Spin(1,3) \ltimes \mathbb{R}^{1,3}$, which is the extension of the Lorentz group $Spin(1,3)$ by translations $\mathbb{R}^{1,3}$ in spacetime, and where the asymmetry of the symbol $\ltimes$ reminds us that the elements of the subgroup $Spin(1,3)$ map $\mathbb{R}^{1,3} \to \mathbb{R}^{1,3}$.

Semidirect sum Some algebras have the structure $\mathfrak{A} = \mathfrak{A}_1 \dotplus \mathfrak{A}_2$, where for every $a, b \in \mathfrak{A}_1$ and $c, d \in \mathfrak{A}_2$ it is true that

$$a * b \in \mathfrak{A}_1, \quad \text{but} \quad c * d,\ a * c,\ c * a \in \mathfrak{A}_2, \tag{B.19}$$

and where $*$ is a "multiplication" in the algebra $\mathfrak{A}$. Formally,

$$\mathfrak{A}_1 * \mathfrak{A}_1 \in \mathfrak{A}_1, \quad \text{but} \quad \mathfrak{A}_1 * \mathfrak{A}_2,\ \mathfrak{A}_2 * \mathfrak{A}_1,\ \mathfrak{A}_2 * \mathfrak{A}_2 \in \mathfrak{A}_2. \tag{B.20}$$

The algebra $\mathfrak{A}$ is said to be a $\mathfrak{A}_2$-extension of the algebra $\mathfrak{A}_1$. The asymmetry symbol "$\dotplus$" here reminds us that $\mathfrak{A}_1$ maps $\mathfrak{A}_1 : \mathfrak{A}_2 \xrightarrow{*} \mathfrak{A}_2$, but it is not a standard notation in the literature, where mostly the uninformative symmetrical symbols $+$ and $\oplus$ are used, and it is left to the Reader to figure out from the context the direction of the inherently asymmetrical relation, i.e., whether $\mathfrak{A}_1 * \mathfrak{A}_2 \in \mathfrak{A}_2$ or $\mathfrak{A}_2 * \mathfrak{A}_1 \in \mathfrak{A}_1$.

Semigroup A collection of elements S equipped with a binary operation $\star$ that satisfies the two axioms [☛ Appendix A.1.1]:
closure $\forall a, b \in S$, $(a \star b) \in S$;
associativity $\forall a, b, c \in S$, $(a \star b) \star c = a \star (b \star c)$.
That is, a semigroup is an associative magma.

Signature In every real n-dimensional vector space V (over the scalar field $\Bbbk$) in which the scalar product $g(v_1, v_2) \in \Bbbk$ is defined for every $v_1, v_2 \in V$, one may find a basis in which $g(\ ,\ )$ is a diagonal matrix. For real vector fields (where $\Bbbk = \mathbb{R}$) the number of positive, negative and vanishing diagonal elements in the diagonalized $g(\ ,\ )$ is called the *signature*. The metric tensor (3.19), $(\eta_{\mu\nu}) = \mathrm{diag}(1, -1, -1, -1)$, in $(3{+}1)$-dimensional spacetime has the signature $(1,3)$. A group of linear transformations is also said to have signature $(1,3)$ if those transformations preserve the scalar product (3.17) defined by the metric tensor of signature $(1,3)$; such transformations form the group $O(1,3)$; $SO(1,3)$ is the subgroup of transformations the determinant of which equals $+1$.

Span A maximal collection of linearly independent elements $\hat{e}_i$, $i = 1, 2, 3 \ldots$, is said to *span* the vector space $V := \{v^i\,\hat{e}_i,\ v^i \in \Bbbk\}$ over a given field of scalars $\Bbbk$.

Spin Intrinsic (albeit perhaps fictitious) angular momentum of an object (particle or physical system) X, meaning that under rotations of the coordinate system the *orientation* of the object X transforms as a representation of the rotation group with the given "angular momentum." For example, a photon has spin $1\hbar$, meaning that its orientation (i.e., polarization) transforms as a spin-$1\hbar$ (vector) representation of the rotation group, the electron as a spin-$\frac{1}{2}\hbar$ (spinor) representation of the rotation group, and the graviton as a spin-$2\hbar$ (rank-2 tensor) representation. The spin of composite systems is the vector sum of all angular momenta of its constituents,[2] but the spin of an elementary particle is not the result of any rotation: elementary particles are *point-like*.

Stückelberg–Feynman interpretation The antiparticle is identified with the particle moving backwards in time. This interpretation follows from the fact that if $\Psi(\mathrm{x})$ is the wave-function of the particle, then its Hermitian conjugate (and, for spin-$\frac{1}{2}$ particles, also the right multiple

[2] The spin of the hydrogen atom as a bound state of an electron and a proton is the vector sum of the orbital angular momentum of the electron in its orbit around the proton, as well as the electron's and the proton's spin.

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

by $\boldsymbol{\gamma}^0$) produces the wave-function of the antiparticle. Expanding into a Fourier series we have that $\Psi(\mathbf{x}) = \sum_\omega e^{i\omega t}\psi_\omega(\vec{r})$, so the Hermitian conjugation is formally identical with the reversal of time.

Surjection A mapping $f : X \twoheadrightarrow Y$, such that for every $y \in Y$ there is an $x \in X$ such that $f(x) = y$.

Symmetry breaking vs. violation A particular process is said to violate a symmetry X if either (1) the X-image of the process does not occur as frequently, i.e., with the same probability, as the original process, or (2) the conserved quantity corresponding to symmetry X is not constant (conserved) during the considered process.

In turn, the symmetry X is broken in a physical system if either (1) the symmetry does not preserve some of the conditions (such as a boundary condition) required of the concrete physical system, or (2) X does not commute with the full Hamiltonian of the system.

Tachyon A particle that propagates through vacuum faster than light, and has an imaginary mass; the appearance of tachyons indicates that the vacuum is not stable [☞ Digression 7.1 on p. 261].

Tangent bundle A vector bundle $\mathscr{T}_{\mathscr{X}} := E(\mathscr{X}; T_{\mathscr{X}}; \pi)$ where $T_x(\mathscr{X}) \cong T_{\mathscr{X}}$ is the tangent space of the space $\mathscr{X}$ at the point $x \in \mathscr{X}$. If x^μ are local coordinates in the space $\mathscr{X}$ at the given point, then $T_x(\mathscr{X})$ may be represented as the vector space of linear combinations $v^\mu \frac{\partial}{\partial x^\mu}$.

Tardion a particle that propagates through vacuum slower than light, and has a real mass; all known matter (and anti-matter) is tardionic, whereupon this term is rarely used.

Tensor product The most general bilinear product of two algebraic structures of the same type, such as vector spaces, algebras, etc. Let X and Y be two vector spaces over the same field, $\Bbbk$. The elements of the tensor product $X \otimes Y$ are $\Bbbk$-linear combinations of elements of the direct product of the sets of elements X and Y, where additionally one requires that the pairs of elements satisfy the relations

$$R := \left\{ \begin{array}{c} e(x{+}x', y) \sim e(x, y) + e(x', y), \quad e(x, y{+}y') \sim e(x, y) + e(x, y'), \\ c\, e(x, y) \sim e(cx, y) \sim e(x, cy). \end{array} \right. \tag{B.21}$$

Then formally,

$$X \otimes Y = \Big\{ \sum_i c_i\, e(x_i, y_i) \ : \ c_i \in \Bbbk,\ (x_i, y_i) \in X \times Y \Big\} \Big/ R, \tag{B.22}$$

which is again a vector space. Similarly, the tensor product of two algebras is again an algebra. In other words, the tensor product inherits the algebraic structure of its factors. Alternatively, Definition B.6 on p. 514 also holds – given using the components with respect to any chosen basis.

Topological space A set of elements ("points") $\mathscr{X}$ with the *topology* τ, which consists of a collection of subsets of the set $\mathscr{X}$ such that they satisfy the axioms:

1. The empty set and the whole set $\mathscr{X}$ belong to τ.
2. The union of an arbitrary number of sets in τ is also in τ.
3. The intersection of an arbitrary finite number of sets in τ is also in τ.

For this system of axioms, the sets in τ are called *open subsets* of the set $\mathscr{X}$; every point $x \in X$ is contained in at least one such open subset, which is then called the *open neighborhood* of the point x. There also exists a complementary definition of topology, using *closed subsets* of the set $\mathscr{X}$; the empty set and the set $\mathscr{X}$ itself here too belong to τ. [☛ also Hausdorff space]

Torsion [☛ curvature]

Vector bundle Let $\mathscr{X}$ be the "base" space, equipped with a copy of a vector space V_x at every point $x \in \mathscr{X}$ of the base space, so that the vector spaces V_x transform homogeneously one into another when the basis point x moves through the base space. The union $\bigcup_{x\in\mathscr{X}} V_x$ is then called the vector bundle over the base space $\mathscr{X}$.

There is also a reverse definition: the total space $E(\mathscr{X};V;\pi)$ of a vector bundle with a given vector space V over the base space $\mathscr{X}$ is such that π is the "vertical" projection with the property that $\pi(E)=\mathscr{X}$, and $\pi^{-1}(x)=V_x\cong V$ for each $x\in\mathscr{X}$.

Vector space A collection of elements (vectors) of which every linear combination with coefficients from a field $\Bbbk$ is also an element of this collection is called a vector space V over the field $\Bbbk$.

Warp, weft and woof are the mutually transversal strands of yarn in a simply woven fabric: *warp* stretches lengthwise from beginning to end, and the strand that is woven left to right and back, weaving through the strands of warp, is variably called *weft* or *woof*.

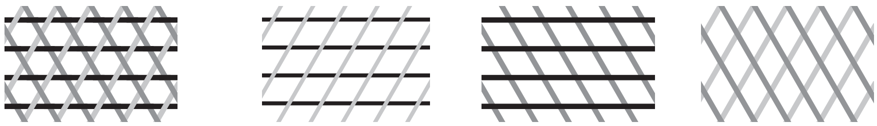

Figure B.1 The triple weave: leaving out any one of the strands dissolves the fabric.

In the theoretical fundamental physics as described herein, the three conceptual strands are provided by (**1**) the Democritean idea of a smallest portion of matter that shows no further, internal constituents, (**2**) the gauge principle of local symmetry, which provides a coherent description of all known fundamental interactions, and (**3**) the idea that all of Nature is to be understood within a unified, comprehensive and logically consistent framework. The (M- and F-theory extended) superstring theoretical system is a framework that conceptually unifies all matter, all of its interactions, as well as the spacetime in which they exist. [☛ Section 1.3.3; Chapters 5–7 and 9]

Wormholes The region in spacetime shaped as a "tunnel," $\mathbb{R}^r\times K^{d-r}$ for $1\leqslant r<d$, where d denotes the total dimension of spacetime, and which either connects two otherwise distant regions of one spacetime, or two otherwise separate spacetimes; K^{d-r} is some compact space (e.g., the 2-sphere, S^2) and represents the "cross-section" of the "tunnel." In known examples, the size of the "cross-section" is typically very small, of the order of $\ell_P\sim10^{-35}$ m and most often has a nonzero size only for a very short time, $t_P\sim10^{-43}$ s. The matter required to keep the wormhole open for a material body or even light to pass through must have "exotic" properties (negative energy density and/or pressure). [☛ Section 9.3.4, Einstein–Rosen bridge]

Yang–Mills interaction, symmetry, theory A gauge interaction, model, symmetry and/or theory is said to be of Yang–Mills type when the gauge 4-vector potential, $\mathbb{A}_\mu\propto(D_\mu-\partial_\mu)$, is the fundamental physical degree of freedom that describes such an interaction. This is the case with electromagnetic, strong and weak nuclear interactions [☞ Chapters 5 and 6], but not with gravity: there, the Christoffel symbol, $\boldsymbol{\Gamma}\propto(D_\mu-\partial_\mu)$, may be expressed as an algebraic combination of the inverse metric tensor and the derivatives of the metric tensor [☞ Chapter 9].

Yukawa field, potential (screened Coulomb field, potential) The Yukawa potential in d-dimensional space is $\Phi_Y=e^{-r/r_0}/r^{d-2}$ and the Yukawa field is $-\vec{\nabla}\Phi_Y$; the negative sign is chosen so that the $r_0\to\infty$ limiting case of the Yukawa field coincides with the traditional definition of the electrostatic field. Here, r_0 is the range of the Yukawa potential and the field.

Yukawa interaction (Yukawa coupling) The coupling between matter field $\Psi(\mathrm{x})$ and the Yukawa potential $\Phi(\mathrm{x})$ produced by the Lagrangian density term $h_\Psi\overline{\Psi}\Phi\Psi$, where h_Ψ is the Yukawa coupling parameter [☛ contact interaction].

ZJBV-quantization [☛ BV-quantization]

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

B.2 Tensor calculus basics

We start with 4-tuples of coordinates such as $\mathrm{x}=(x^0,x^1,x^2,x^3)$, two functions of such coordinates, f and g, and the well-known derivative rules in multi-variate calculus:

$$\textbf{product rule}\quad \frac{\partial}{\partial x^\mu}\big(f(\mathrm{x})\,g(\mathrm{x})\big)=\Big(\frac{\partial f(\mathrm{x})}{\partial x^\mu}\Big)g(\mathrm{x})+f(\mathrm{x})\Big(\frac{\partial g(\mathrm{x})}{\partial x^\mu}\Big),\tag{B.23}$$

$$\textbf{chain rule}\quad \frac{\partial}{\partial x^\mu}\Big(y^\nu\big(\mathrm{z}(\mathrm{x})\big)\Big)=\Big(\frac{\partial y^\nu}{\partial z^\rho}\Big)\Big(\frac{\partial z^\rho}{\partial x^\mu}\Big).\tag{B.24}$$

Taking $\mathrm{x}=(x^0,x^1,x^2,x^3)$, $\mathrm{y}=(y^0,y^1,y^2,y^3)$ and $\mathrm{z}=(z^0,z^1,z^2,z^3)$ to provide *general* coordinate systems, these 4-tuples need not span vector spaces in general: In general coordinate systems, linear combinations $c_\mu x^\mu$ with numerical (dimensionless) constants c_μ need make no sense at all. At the very least, the constants c_μ could be equipped with appropriate physical units. For example, in the familiar spherical coordinate system (r,θ,ϕ), a linear combination such as $(\frac{\pi}{2}\,r-\sqrt{3}\theta)$ makes no sense since the two summands have wholly different physical units. In turn, denoting by L some suitable and constant length, the linear combination $(\frac{\pi}{2L}\,r-\sqrt{3}\theta)$ does make sense in general, although it does not seem to provide any physically reasonable quantity. Even so, and owing to the generally curvilinear nature of general coordinates and their diverse behavior (e.g., $\theta\simeq\theta\pm2\pi$ while $r\geqslant0$), linear combinations (even if adjusted for physical units) of general coordinates do not, in general, represent a point in the space parametrized by these coordinates.

However, owing to the infinitesimal nature of the differentials $\mathrm{d}x^\mu$ and the operators $\frac{\partial}{\partial x^\mu}$, the 4-tuples $(\mathrm{d}x^0,\mathrm{d}x^1,\mathrm{d}x^2,\mathrm{d}x^3)$ and $(\frac{\partial}{\partial x^0},\frac{\partial}{\partial x^1},\frac{\partial}{\partial x^2},\frac{\partial}{\partial x^3})$ *do* span two vector spaces – again with the proviso that the constants in the respective linear combinations may have to be equipped with adequate physical units. The application of the chain rule to these clearly distinguishes them and permits the definition of two *distinct* types of 4-vectors:

$$\textbf{contravariant vector (3.11c)}\quad \mathrm{d}x^\mu=\mathrm{d}y^\nu\Big(\frac{\partial x^\mu}{\partial y^\nu}\Big)\quad\leftrightarrow\quad A^\mu(\mathrm{x})=A^\nu(\mathrm{y})\Big(\frac{\partial x^\mu}{\partial y^\nu}\Big);\tag{B.25}$$

$$\textbf{covariant vector (3.11d)}\quad \frac{\partial}{\partial x^\mu}=\Big(\frac{\partial y^\nu}{\partial x^\mu}\Big)\frac{\partial}{\partial y^\nu}\quad\leftrightarrow\quad B_\mu(\mathrm{x})=\Big(\frac{\partial y^\nu}{\partial x^\mu}\Big)B_\nu(\mathrm{y}),\tag{B.26}$$

simply by observing that they transform with the *opposite* partial derivatives, as was already done in Digression 3.2 on p. 88.

B.2.1 Basis elements

We then proceed as follows: Given any coordinate system $\mathrm{x}:=(x^0,x^1,x^2,x^3)$ equipped with a metric tensor, $g_{\mu\nu}(\mathrm{x})$, we specify:

1. The line element $\mathrm{d}s$ provides the invariant norm of the coordinate differentials:

$$\mathrm{d}s:=\sqrt{\mathrm{dx}\cdot\mathrm{dx}},\quad \mathrm{dx}\cdot\mathrm{dx}:=g_{\mu\nu}(\mathrm{x})\,\mathrm{d}x^\mu\mathrm{d}x^\nu.\tag{B.27}$$

2. The invariant Kronecker symbol

$$\delta^\mu_\nu:=\frac{\partial x^\mu}{\partial x^\nu}=\begin{cases}1 \text{ if } \mu=\nu,\\ 0 \text{ if } \mu\neq\nu,\end{cases}\tag{B.28}$$

is simply the statement that the coordinates x^μ are mutually independent.

3. The invariant Levi-Civita symbol is defined implicitly by expanding the Jacobian of a coordinate transformation $\mathrm{x}\to\mathrm{y}$:

$$\Big|\frac{\partial\mathrm{x}}{\partial\mathrm{y}}\Big|=:\varepsilon^{\mu\nu\rho\sigma}\frac{\partial x^0}{\partial y^\mu}\frac{\partial x^1}{\partial y^\nu}\frac{\partial x^2}{\partial y^\rho}\frac{\partial x^3}{\partial y^\sigma}=:\varepsilon_{\mu\nu\rho\sigma}\frac{\partial x^\mu}{\partial y^0}\frac{\partial x^\nu}{\partial y^1}\frac{\partial x^\rho}{\partial y^2}\frac{\partial x^\sigma}{\partial y^3},\tag{B.29}$$

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

that is,

$$\varepsilon^{\mu\nu\rho\sigma} = \varepsilon_{\mu\nu\rho\sigma} := \begin{cases} +1 & \text{if } \mu,\nu,\rho,\sigma = \text{even permutation of } 0,1,2,3; \\ -1 & \text{if } \mu,\nu,\rho,\sigma = \text{odd permutation of } 0,1,2,3; \\ 0 & \text{otherwise.} \end{cases} \tag{B.30}$$

4. It follows that

$$\varepsilon_{\alpha\beta\gamma\delta}\varepsilon^{\mu\nu\rho\sigma} = 4!\,\delta^{\mu\nu\rho\sigma}_{[\alpha\beta\gamma\delta]}, \tag{B.31}$$

where

$$\begin{gathered} \delta^{\mu\nu}_{[\alpha\beta]} := \tfrac{1}{2}\big(\delta^{\mu}_{\alpha}\delta^{\nu}_{\beta} - \delta^{\mu}_{\beta}\delta^{\nu}_{\alpha}\big), \quad \delta^{\mu\nu\rho}_{[\alpha\beta\gamma]} := \tfrac{1}{3}\big(\delta^{\mu\nu}_{[\alpha\beta]}\delta^{\rho}_{\gamma} + \delta^{\mu\nu}_{[\beta\gamma]}\delta^{\rho}_{\alpha} + \delta^{\mu\nu}_{[\gamma\alpha]}\delta^{\rho}_{\beta}\big), \\ \text{and so } \delta^{\mu\nu\rho\sigma}_{[\alpha\beta\gamma\delta]} := \tfrac{1}{4}\big(\delta^{\mu\nu\rho}_{[\alpha\beta\gamma]}\delta^{\sigma}_{\delta} - \delta^{\mu\nu\rho}_{[\delta\alpha\beta]}\delta^{\sigma}_{\gamma} + \delta^{\mu\nu\rho}_{[\gamma\delta\alpha]}\delta^{\sigma}_{\beta} - \delta^{\mu\nu\rho}_{[\beta\gamma\delta]}\delta^{\sigma}_{\alpha}\big). \end{gathered} \tag{B.32}$$

5. Owing to the reciprocal transformation rules (3.11c)–(3.11d), the *contractions*

$$\mathrm{A}(\mathrm{x})\cdot\mathrm{B}(\mathrm{x}) = A^{\mu}(\mathrm{x})\,B_{\mu}(\mathrm{x}), \quad \mathrm{A}(\mathrm{x})\cdot\partial = A^{\mu}(\mathrm{x})\partial_{\mu}, \quad \mathrm{dx}\cdot\mathrm{B}(\mathrm{x}) = \mathrm{d}x^{\mu}\,B_{\mu}(\mathrm{x}), \tag{3.12a$'$}$$

$$\text{and} \qquad \mathrm{d} := \mathrm{dx}\cdot\partial := \mathrm{d}x^{\mu}\frac{\partial}{\partial x^{\mu}} \tag{B.33}$$

are all invariant under general coordinate transformations $x^{\mu} \mapsto y^{\mu}(\mathrm{x})$, as specified in Definition 9.1 on p. 319. Thus, the $\mathrm{d}x^{\mu}$ may be used as basis vectors for covariant components $B_{\mu}(\mathrm{x})$, and the ∂_{μ} may be used as basis vectors for contravariant components $A^{\mu}(\mathrm{x})$. This is the typical *choice* in the mathematics literature as it connects tensor algebra and differential geometry; see Comment B.1 on p. 512.

6. Let $\mathrm{e}(\mathrm{x})$ denote an event – a point in spacetime specified, with the coordinates x, and let the displacement to an infinitesimally near event be $\mathrm{de} = \frac{\partial \mathrm{e}}{\partial x^{\mu}}\mathrm{d}x^{\mu}$, expressed in the x^{μ} coordinates. Then, we define:

$$\textbf{covariant basis element} \quad \mathrm{e}_{\mu}(\mathrm{x}) := \frac{\partial \mathrm{e}}{\partial x^{\mu}}, \tag{B.34}$$

$$\textbf{contravariant basis element} \quad \mathrm{e}^{\mu}(\mathrm{x}) := g^{\mu\nu}\mathrm{e}_{\nu}(\mathrm{x}). \tag{B.35}$$

The scalar product of these basis elements is defined so that

$$\mathrm{e}_{\mu}(\mathrm{x})\cdot\mathrm{e}_{\nu}(\mathrm{x}) = g_{\mu\nu}(\mathrm{x}), \quad \mathrm{e}^{\mu}(\mathrm{x})\cdot\mathrm{e}^{\nu}(\mathrm{x}) = g^{\mu\nu}(\mathrm{x}) \quad \text{and} \quad \mathrm{e}_{\mu}(\mathrm{x})\cdot\mathrm{e}^{\nu}(\mathrm{x}) = \delta^{\nu}_{\mu}. \tag{B.36}$$

7. Given the contravariant components of a 4-vector, $A^{\mu}(\mathrm{x})$, the 4-vector is invariantly specified as $\mathrm{A}(\mathrm{x}) = A^{\mu}(\mathrm{x})\,\mathrm{e}_{\mu}(\mathrm{x})$. Given the covariant components of a 4-vector, $B_{\mu}(\mathrm{x})$, the 4-vector is invariantly specified as $\mathrm{B}(\mathrm{x}) = B_{\mu}(\mathrm{x})\,\mathrm{e}^{\mu}(\mathrm{x})$.

Here, "invariant," "covariant" and "contravariant" all refer to transformation properties with respect to the general coordinate transformations specified in Definition 9.1 on p. 319.

Given the definition of contravariant vectors (B.25), it is straightforward to compute the transformation rule for the differential "volume" element:

$$\begin{aligned} \mathrm{d}^4x = \mathrm{d}x^0\mathrm{d}x^1\mathrm{d}x^2\mathrm{d}x^3 &= \tfrac{1}{4!}\varepsilon_{\mu\nu\rho\sigma}\,\mathrm{d}x^{\mu}\mathrm{d}x^{\nu}\mathrm{d}x^{\rho}\mathrm{d}x^{\sigma} \\ &= \tfrac{1}{4!}\varepsilon_{\mu\nu\rho\sigma}\Big(\frac{\partial x^{\mu}}{\partial y^{\alpha}}\mathrm{d}y^{\alpha}\Big)\Big(\frac{\partial x^{\nu}}{\partial y^{\beta}}\mathrm{d}y^{\beta}\Big)\Big(\frac{\partial x^{\rho}}{\partial y^{\gamma}}\mathrm{d}y^{\gamma}\Big)\Big(\frac{\partial x^{\sigma}}{\partial y^{\delta}}\mathrm{d}y^{\delta}\Big) \\ &= \tfrac{1}{4!}\varepsilon_{\mu\nu\rho\sigma}\frac{\partial x^{\mu}}{\partial y^{\alpha}}\frac{\partial x^{\nu}}{\partial y^{\beta}}\frac{\partial x^{\rho}}{\partial y^{\gamma}}\frac{\partial x^{\sigma}}{\partial y^{\delta}}\,\mathrm{d}y^{\alpha}\mathrm{d}y^{\beta}\mathrm{d}y^{\gamma}\mathrm{d}y^{\delta} \\ &= \tfrac{1}{4!}\varepsilon_{\mu\nu\rho\sigma}\frac{\partial x^{\mu}}{\partial y^{\alpha}}\frac{\partial x^{\nu}}{\partial y^{\beta}}\frac{\partial x^{\rho}}{\partial y^{\gamma}}\frac{\partial x^{\sigma}}{\partial y^{\delta}}\,\delta^{\alpha\beta\gamma\delta}_{[\epsilon\varphi\lambda\kappa]}\,\mathrm{d}y^{\epsilon}\mathrm{d}y^{\varphi}\mathrm{d}y^{\lambda}\mathrm{d}y^{\kappa} \end{aligned} \tag{B.37a}$$

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

$$= \Big[\tfrac{1}{4!}\varepsilon_{\mu\nu\rho\sigma}\,\frac{\partial x^\mu}{\partial y^\alpha}\frac{\partial x^\nu}{\partial y^\beta}\frac{\partial x^\rho}{\partial y^\gamma}\frac{\partial x^\sigma}{\partial y^\delta}\,\varepsilon^{\alpha\beta\gamma\delta}\Big]\,\tfrac{1}{4!}\varepsilon_{\epsilon\varphi\lambda\kappa}\,\mathrm{d}y^\epsilon\mathrm{d}y^\varphi\mathrm{d}y^\lambda\mathrm{d}y^\kappa$$

$$= \det\Big[\frac{\partial \mathrm{x}}{\partial \mathrm{y}}\Big]\,\mathrm{d}^4y, \tag{B.37b}$$

where the key relation (B.31) between the Levi-Civita and the Kronecker symbols was used. We have also used the general expressions for the determinants of $n\times n$ matrices representing rank-2 tensor densities:

$$\textbf{type } (1,1) \qquad \det[\mathbb{M}] := \tfrac{1}{4!}\varepsilon_{\mu_1\cdots\mu_n}\,M^{\mu_1}_{\nu_1}\cdots M^{\mu_n}_{\nu_n}\,\varepsilon^{\nu_1\cdots\nu_n}, \tag{B.38a}$$

$$\textbf{type } (0,2) \qquad \det[\mathbb{N}] := \tfrac{1}{4!}\varepsilon^{\mu_1\cdots\mu_n}\,N_{\mu_1\nu_1}\cdots N_{\mu_n\nu_n}\,\varepsilon^{\nu_1\cdots\nu_n}, \tag{B.38b}$$

$$\textbf{type } (2,0) \qquad \det[\mathbb{P}] := \tfrac{1}{4!}\varepsilon_{\mu_1\cdots\mu_n}\,P^{\mu_1\nu_1}\cdots P_{\mu_n\nu_n}\,\varepsilon_{\nu_1\cdots\nu_n}. \tag{B.38c}$$

Given the definitions of the “ingredients”:

1. a contravariant vector (3.11c),
2. a covariant vector (3.11d),
3. a scalar density (9.8),

we adapt Weyl’s Construction A.1 and generate representations of the group of general coordinate transformations, by taking tensor products of the “ingredients” and symmetrizing like factors in all possible ways. More precisely,

Definition B.1 *Tensor densities may be formally constructed from a scalar density U, a contravariant vector $V = V^\mu \mathrm{e}_\mu$ and a covariant vector $W = W_\mu \mathrm{e}_\mu$:*

$$U(\mathrm{y}) = \Big(\det\Big[\frac{\partial \mathrm{y}}{\partial \mathrm{x}}\Big]\Big)U(\mathrm{x}), \quad V^\mu(\mathrm{y}) = \frac{\partial y^\mu}{\partial x^\nu}\,V^\nu(\mathrm{x}), \quad W_\mu(\mathrm{y}) = \frac{\partial x^\nu}{\partial y^\mu}\,W_\nu(\mathrm{x}). \tag{B.39}$$

One constructs first the vector space of ordered products,

$$T(p,q;w) := U^w\cdot\underbrace{V\otimes\cdots\otimes V}_{p}\otimes\underbrace{W\otimes\cdots\otimes W}_{q}, \tag{B.40}$$

on which the permutation group $S_p\times S_q$ acts, where S_p permutes the V-factors and S_q permutes the W-factors. The vector space $T(p,q;w)$ may then be decomposed, in a unique fashion, into a direct sum of irreducible representations of the permutation group (index symmetrization). Finally, each summand in the so-obtained direct sum may be further decomposed by ***contracting*** *with invariant tensors δ^μ_ν, $\varepsilon_{\mu\nu\rho\sigma}$ and $\varepsilon^{\mu\nu\rho\sigma}$.*

Focusing on the structure of the transformation properties, i.e., how a quantity transforms with respect to general coordinate transformations, rather than how it may have been constructed, produces the complementary general definition:

Definition B.2 (tensor density) *A quantity that is in some coordinate system (with coordinates x^μ) specified by its* ***components*** *$\{T^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}(\mathrm{x})\}$ and the components of which in some other coordinate system (with coordinates y^μ) may be computed using the relations*

$$T^{\rho_1\cdots\rho_p}_{\sigma_1\cdots\sigma_q}(\mathrm{y}) = \Big(\det\Big[\frac{\partial \mathrm{y}}{\partial \mathrm{x}}\Big]\Big)^w\,\frac{\partial y^{\rho_1}}{\partial x^{\mu_1}}\cdots\frac{\partial y^{\rho_p}}{\partial x^{\mu_p}}\,\frac{\partial x^{\nu_1}}{\partial y^{\sigma_1}}\cdots\frac{\partial x^{\nu_q}}{\partial y^{\sigma_q}}\,T^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}(\mathrm{x}) \tag{B.41}$$

is called a ***tensor density*** *of weight w, type (p,q) and rank $p{+}q$. Weight-0 tensor densities are called* ***tensors****; rank-1 tensors are called* ***vectors****, and rank-0 tensors are* ***scalars****, i.e.,* ***invariants****. The symbol $\big[\frac{\partial \mathrm{y}}{\partial \mathrm{x}}\big]$ denotes the matrix of partial derivatives that appear in equation (B.26).*

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

This special meaning of the word "density" – which in this special use always follows the adjective "tensor," "vector" or "scalar" – must not be confused with the familiar notion as in "per unit of volume." Thus, for example, "Lagrangian density" literally means "Lagrangian per unit of volume." On the other hand, in the sense of Definition B.2 and in the typical practice in theoretical and mathematical physics, Lagrangian densities are – as well as the Lagrangians and Hamiltonians and Hamiltonian densities – **scalars**, i.e., weight-0 scalar densities [☞ Conclusion 9.5 on p. 328].

> ***Comment B.1*** *Specifying all components in any one concrete basis* ***does*** *specify the tensor density abstractly, since the relations (B.41) provide the transformation rules from one basis into any other one. In the mathematical literature one typically uses the natural basis* $\{dx^\mu, \frac{\partial}{\partial x^\mu}\}$*, whereby a tensor density is specified* ***invariantly*** *as*
>
> $$T(\mathrm{x}) \;:\quad dx^{\nu_1}\otimes\cdots\otimes dx^{\nu_p}\, T^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}(\mathrm{x})\,\frac{\partial}{\partial x^{\mu_1}}\otimes\cdots\otimes\frac{\partial}{\partial x^{\mu_q}}. \tag{B.42a}$$
>
> *Using the relations (B.25)–(B.26) and (B.41), it is then easy to show that*
>
> $$T(\mathrm{y}) = \Big(\det\Big[\frac{\partial \mathrm{y}}{\partial \mathrm{x}}\Big]\Big)^w\, T(\mathrm{x}). \tag{B.42b}$$
>
> *In this book, I follow the physicists' practice of specifying and manipulating components (with respect to any one particular basis) as the representatives of the whole tensor density; see Digression 3.3 on p. 88, as well as the discussion in Wald's textbook [548].*

B.2.2 Tensor algebra

Scalar functions (weight-0 scalar densities) over spacetime are, in the physics nomenclature, typically called scalar fields. These scalar fields (in the physics sense) form – at every spacetime point separately – a field in the mathematical sense. That is, addition and multiplication of scalar fields – taken at any particular spacetime point – follows the usual rules of addition and multiplication of "ordinary" (real and complex) numbers. It is, however, important to note that this is not the case when adding/multiplying scalar fields where the summands/factors are taken at different spacetime points: $f(\mathrm{x})\,g(\mathrm{y})$ is not a function of either just x or just y, but of both. Thus, scalar functions (over the whole spacetime) do not form the usual algebraic structure of a field. However, restricting the binary operations to the cases when both summands/factors are taken at the same spacetime point produces an algebraic structure that minimally deviates from the standard definition of the (mathematical) field, i.e., extends this definition.[3] The corresponding generalizations of functions (and all the tensor densities as well) over general, curved spaces are called *sections* of various *bundles* [☞ [563, 210, 379, 176], to begin with].

Similarly, tensor densities $T^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}(\mathrm{x})$ may be multiplied by scalar densities $f(\mathrm{x})$ by simply multiplying each component. Also, it should be clear that the tensor densities of the same type and weight may be added, which permits defining point-by-point linear combinations such as

$$f(\mathrm{x})\,T^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}(\mathrm{x}) + h(\mathrm{x})\,U^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}(\mathrm{x}), \tag{B.43}$$

as long as the sum of weights of f and T equals the sum of weights of h and U, and this generates a structure that minimally generalizes the structure of a vector space:

[3] The deviation pertains precisely to the general case, when the arguments of the two factors in a product are not the same. For those cases, one may simply declare that multiplication is not defined – which is already a departure from the standard definition of a field, or one may define such a product via some formal expansion into a series in powers of the difference $(\mathrm{x}-\mathrm{y})$ – when such a power series is well defined, etc.

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

Definition B.3 *Tensor densities of the same type form a generalization of the vector space as their linear combination is defined by specifying*

$$f(\mathrm{x})\,T^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}(\mathrm{x}) + h(\mathrm{x})\,U^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}(\mathrm{x}), \tag{B.44}$$

where the coefficients are scalar densities of complementary weights:

$$w\big[\,f(\mathrm{x})\,T^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}(\mathrm{x})\,\big] = w\big[\,h(\mathrm{x})\,U^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}(\mathrm{x})\,\big]. \tag{B.45}$$

The linearity of the definition guarantees that the result (B.44) is again a tensor density of the same rank, type and weight.

The structure of a vector space is recovered by restricting to constant coefficients and tensor densities of the same weight.

The following two operations are also important:

Definition B.4 (Contraction) *For any type-(p,q) tensor density, where $p \neq 0 \neq q$, one constructs the* ***contraction***

$$\delta^{\nu_i}_{\mu_j} \;:\; T^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}(\mathrm{x}) \mapsto T^{\mu_1\cdots\widehat{\mu}_j\cdots\mu_p}_{\nu_1\cdots\widehat{\nu}_i\cdots\nu_q}(\mathrm{x}) = \Big(\delta^{\nu_i}_{\mu_j}\,T^{\mu_1\cdots\mu_j\cdots\mu_p}_{\nu_1\cdots\nu_i\cdots\nu_q}(\mathrm{x})\Big), \tag{B.46}$$

where $\widehat{\mu}_i$ denotes that the index μ_i is omitted from the sequence. The result of contracting is a type-$(p-1,q-1)$ tensor density of the same weight as the original tensor density.

Definition B.5 *For any two indices of the same type, one defines*

$$T^{(\mu\nu)\cdots}_{\cdots} := \tfrac{1}{2}\big(T^{\mu\nu\cdots}_{\cdots} + T^{\nu\mu\cdots}_{\cdots}\big), \quad \textit{and} \quad T^{[\mu\nu]\cdots}_{\cdots} := \tfrac{1}{2}\big(T^{\mu\nu\cdots}_{\cdots} - T^{\nu\mu\cdots}_{\cdots}\big), \tag{B.47}$$

the so-called symmetric and antisymmetric part of the original tensor density. The linearity of the definition guarantees that both parts retain the rank, type and weight of the original tensor density.

With tensor densities of a rank higher than two, the combinatorial possibilities and wealth of various (anti)symmetrization patterns grow very quickly; some simple examples are given in relations (A.66) and (A.76). Technically more precisely, the various forms of (anti)symmetrization provide various representations of the permutation group that acts by permuting the indices of the same type (here, subscript vs. superscripts).

Comment B.2 *Every tensor density with at least two indices of the same type may always be decomposed:*

$$T^{\mu\nu\cdots}_{\cdots} \equiv 2\cdot\tfrac{1}{2}T^{\mu\nu\cdots}_{\cdots} + \tfrac{1}{2}T^{\nu\mu\cdots}_{\cdots} - \tfrac{1}{2}T^{\nu\mu\cdots}_{\cdots} = T^{(\mu\nu)\cdots}_{\cdots} + T^{[\mu\nu]\cdots}_{\cdots}, \tag{B.48}$$

where $T^{(\mu\nu)\cdots}_{\cdots}$ and $T^{[\mu\nu]\cdots}_{\cdots}$ transform the same as the original tensor density, $T^{\mu\nu\cdots}_{\cdots}$. More generally, every tensor density may be decomposed into a sum of tensor densities, each of which is an irreducible representation of the permutation group that acts by permuting indices of the same type.

The operations provided by the definitions B.3, B.4 and B.5 generate a structure that is usually called simply "linear algebra."

Finally, define also the multiplication of tensor densities:

Definition B.6 *For any two tensor densities $T^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}(\mathrm{x})$ and $U^{\rho_1\cdots\rho_{p'}}_{\sigma_1\cdots\sigma_{q'}}(\mathrm{x})$, respectively of type (p,q) and $(p'q')$ and weights w and w', the* ***tensor product*** *may be specified by the relation*

$$(T\otimes U)^{\mu_1\cdots\mu_{p+p'}}_{\nu_1\cdots\nu_{q+q'}}(\mathrm{x}) := T^{\mu_1\cdots\mu_p}_{\nu_1\cdots\nu_q}(\mathrm{x})\, U^{\mu_{p+1}\cdots\mu_{p+p'}}_{\nu_{q+1}\cdots\nu_{q+q'}}(\mathrm{x}) \tag{B.49}$$

the result of which is a type-$(p+p',q+q')$ and weight-$(w+w')$ tensor density.

B.2.3 Tensor calculus

The *rate of change* of a vector such as $\mathrm{A}(\mathrm{x}) = A^\mu(\mathrm{x})\,\mathrm{e}_\mu(\mathrm{x})$ over spacetime is then

$$\frac{\partial \mathrm{A}}{\partial x^\mu} = \frac{\partial}{\partial x^\mu}\Big(A^\nu(\mathrm{x})\,\mathrm{e}_\nu(\mathrm{x})\Big) = \frac{\partial A^\nu}{\partial x^\mu}\,\mathrm{e}_\nu(\mathrm{x}) + A^\nu(\mathrm{x})\,\frac{\partial \mathrm{e}_\nu}{\partial x^\mu}. \tag{B.50}$$

Since e_ν form a complete set, the partial derivative in the second term must be expressible as a linear combination in the same basis:

$$\frac{\partial \mathrm{e}_\nu}{\partial x^\mu} =: \Gamma^\rho_{\mu\nu}(\mathrm{x})\,\mathrm{e}_\rho(\mathrm{x}), \tag{B.51}$$

where $\Gamma^\rho_{\mu\nu}(\mathrm{x})$ are, for each pair (μ,ν) and at each point x in spacetime, simply the 4-tuple of coefficient functions in the linear combination of basis vectors $\mathrm{e}_\rho(\mathrm{x})$. Combining results (B.50) and (B.51), we have

$$\frac{\partial \mathrm{A}}{\partial x^\mu} = \Big[\frac{\partial A^\rho}{\partial x^\mu} + A^\nu\,\Gamma^\rho_{\mu\nu}\Big]\mathrm{e}_\rho(\mathrm{x}). \tag{B.52}$$

It is straightforward that

$$\frac{\partial}{\partial x^\mu}\Big(\mathrm{e}_\nu{\cdot}\mathrm{e}^\rho = \delta^\rho_\nu\Big) = 0 \quad\Rightarrow\quad \frac{\partial \mathrm{e}^\rho}{\partial x^\mu} = -\Gamma^\rho_{\mu\nu}(\mathrm{x})\,\mathrm{e}^\nu(\mathrm{x}), \tag{B.53}$$

whereby

$$\frac{\partial \mathrm{B}}{\partial x^\mu} = \frac{\partial B_\nu}{\partial x^\mu}\,\mathrm{e}^\nu(\mathrm{x}) + B_\nu(\mathrm{x})\,\frac{\partial \mathrm{e}^\nu}{\partial x^\mu} = \Big[\frac{\partial B_\nu}{\partial x^\mu} - B_\rho\,\Gamma^\rho_{\mu\nu}\Big]\mathrm{e}^\nu(\mathrm{x}). \tag{B.54}$$

The quantities in the square brackets in equations (B.52) and (B.54) are then defined as the *covariant derivatives* of the components

$$D_\mu A^\rho := \big[\partial_\mu A^\rho + \Gamma^\rho_{\mu\nu}A^\nu\big] \quad\text{and}\quad D_\mu B_\nu := \big[\partial_\mu B_\nu - \Gamma^\rho_{\mu\nu}B_\rho\big]. \tag{B.55}$$

The formula (9.17) is then the straightforward iteration of these two definitions, as dictated by Weyl's Construction A.1 on p. 478, adapted here to provide Definition B.1 on p. 511.

The definition of $\Gamma^\rho_{\mu\nu}(\mathrm{x})$ in equation (B.51) and the relations (B.36) then imply several important properties of $\Gamma^\rho_{\mu\nu}(\mathrm{x})$. First,

$$\Gamma^\rho_{\mu\nu}\,\mathrm{e}_\rho = \frac{\partial\,\mathrm{e}_\nu}{\partial x^\mu} = \frac{\partial^2\,\mathrm{e}}{\partial x^\mu\partial x^\nu} = \frac{\partial^2\,\mathrm{e}}{\partial x^\nu\partial x^\mu} = \Gamma^\rho_{\nu\mu}\,\mathrm{e}_\rho \quad\Rightarrow\quad \Gamma^\rho_{\mu\nu} = \Gamma^\rho_{\nu\mu}. \tag{B.56}$$

Next, compute

$$\frac{\partial g_{\mu\nu}}{\partial x^\rho} = \frac{\partial}{\partial x^\rho}(\mathrm{e}_\mu{\cdot}\mathrm{e}_\nu) = \Gamma^\sigma_{\mu\rho}\mathrm{e}_\sigma{\cdot}\mathrm{e}_\nu + \mathrm{e}_\mu{\cdot}\Gamma^\sigma_{\nu\rho}\mathrm{e}_\sigma = \Gamma^\sigma_{\mu\rho}g_{\sigma\nu} + g_{\mu\sigma}\Gamma^\sigma_{\nu\rho}. \tag{B.57}$$

Reusing this equality with permuted indices μ,ν,ρ, we obtain

$$\frac{\partial g_{\mu\sigma}}{\partial x^\nu} + \frac{\partial g_{\nu\sigma}}{\partial x^\mu} - \frac{\partial g_{\mu\nu}}{\partial x^\sigma} = 2g_{\sigma\rho}\Gamma^\rho_{\mu\nu}, \tag{B.58}$$

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

which implies the standard formula [508, 62, 367, 548, 66, 96]

$$\Gamma^{\rho}_{\mu\nu} = \tfrac{1}{2} g^{\rho\sigma} \Big[\frac{\partial g_{\mu\sigma}}{\partial x^{\nu}} + \frac{\partial g_{\nu\sigma}}{\partial x^{\mu}} - \frac{\partial g_{\mu\nu}}{\partial x^{\sigma}} \Big]. \tag{B.59}$$

It is then straightforward to show that

$$D_{\mu} g_{\nu\rho} = 0 = D_{\mu} g^{\nu\rho}. \tag{B.60}$$

We close with a useful result and a comment. The Jacobi identity for derivatives of the determinant $g := \det[g_{..}]$ is

$$\frac{\partial g}{\partial x^{\mu}} = g\, g^{\nu\rho} \frac{\partial g_{\nu\rho}}{\partial x^{\mu}} \quad \Rightarrow \quad g^{\nu\rho} \frac{\partial g_{\nu\rho}}{\partial x^{\mu}} = \frac{1}{g} \frac{\partial g}{\partial x^{\mu}} = \frac{1}{(-g)} \frac{\partial(-g)}{\partial x^{\mu}} = \frac{\partial \ln(-g)}{\partial x^{\mu}}, \tag{B.61}$$

where the sign-change was necessary as spacetime metrics have an odd number of negative eigenvalues and so a negative determinant. Now contract the expression (B.59):

$$\Gamma^{\mu}_{\mu\nu} = \tfrac{1}{2} g^{\mu\sigma} \Big[\frac{\partial g_{\mu\sigma}}{\partial x^{\nu}} + \frac{\partial g_{\nu\sigma}}{\partial x^{\mu}} - \frac{\partial g_{\mu\nu}}{\partial x^{\sigma}} \Big] = \tfrac{1}{2} g^{\mu\sigma} \frac{\partial g_{\mu\sigma}}{\partial x^{\nu}}, \tag{B.62}$$

since $g^{\mu\sigma} \frac{\partial g_{\nu\sigma}}{\partial x^{\mu}} \overset{\mu:\sigma}{=} g^{\sigma\mu} \frac{\partial g_{\nu\mu}}{\partial x^{\sigma}} = g^{\mu\sigma} \frac{\partial g_{\mu\nu}}{\partial x^{\sigma}}$ and the last two terms cancel. Using then the identity (B.61) yields

$$\Gamma^{\mu}_{\mu\nu} = \tfrac{1}{2} \frac{\partial \ln(g)}{\partial x^{\mu}} = \frac{\partial \ln\big(\sqrt{g}\big)}{\partial x^{\mu}} = \frac{1}{\sqrt{g}} \frac{\partial \sqrt{g}}{\partial x^{\mu}}. \tag{B.63}$$

Therefore,

$$\begin{aligned} D\cdot \mathrm{A} &= \mathrm{e}^{\mu} \cdot \frac{\partial \mathrm{A}}{\partial x^{\mu}} = (D_{\mu} A^{\nu})\, \mathrm{e}^{\mu} \cdot \mathrm{e}_{\nu} = (D_{\mu} A^{\mu}) = \frac{\partial A^{\mu}}{\partial x^{\mu}} + \Gamma^{\mu}_{\mu\nu} A^{\nu}, \\ &= \frac{\partial A^{\nu}}{\partial x^{\nu}} + \Big(\frac{1}{\sqrt{g}} \frac{\partial \sqrt{g}}{\partial x^{\nu}} \Big) A^{\nu} = \frac{1}{\sqrt{g}} \frac{\partial(\sqrt{g} A^{\nu})}{\partial x^{\nu}} = \frac{1}{\sqrt{g}} \frac{\partial(\sqrt{g}\, g^{\nu\rho} A_{\rho})}{\partial x^{\nu}} \end{aligned} \tag{B.64}$$

provides the definition of the spacetime gradient of a 4-vector, alternatively given for a vector specified in terms of contravariant and covariant components. The spacetime gradient of a type-(p,q) tensor density of weight w is then obtained by iterating this result. For example, the spacetime divergence of a type-$(2,0)$ tensor is

$$(D\cdot \mathbb{T})^{\nu} = \frac{\partial T^{\mu\nu}}{\partial x^{\mu}} + \Gamma^{\mu}_{\mu\sigma} T^{\sigma\nu} + \Gamma^{\nu}_{\mu\sigma} T^{\mu\sigma} = \frac{1}{\sqrt{g}} \frac{\partial\big(\sqrt{g} T^{\sigma\nu}\big)}{\partial x^{\sigma}} + \Gamma^{\nu}_{\mu\sigma} T^{\mu\sigma}. \tag{B.65}$$

The general result is

$$\big(\widetilde{D_{\lambda}\, \mathbb{T}}(\mathrm{y})\big)^{\rho_1 \cdots \rho_p}_{\sigma_1 \cdots \sigma_q} = \Big(\det \Big[\frac{\partial \mathrm{y}}{\partial \mathrm{x}} \Big| \Big)^{w} \frac{\partial y^{\rho_1}}{\partial x^{\mu_1}} \cdots \frac{\partial y^{\rho_p}}{\partial x^{\mu_p}} \frac{\partial x^{\nu_1}}{\partial y^{\sigma_1}} \cdots \frac{\partial x^{\nu_q}}{\partial y^{\sigma_q}} \frac{\partial x^{\kappa}}{\partial y^{\lambda}} \big(D_{\kappa}\, \mathbb{T}(\mathrm{x})\big)^{\mu_1 \cdots \mu_p}_{\nu_1 \cdots \nu_q}. \tag{B.66}$$

That is, the covariant derivative of a type-(p,q) tensor density of weight w is a type-$(p,q{+}1)$ tensor density of weight w.

Finally, we note that for every μ the vector $\mathrm{e}_{\mu}(\mathrm{x})$ is defined infinitesimally near the point x. Using the 4-vector of partial derivatives $\frac{\partial}{\partial x^{\nu}}$, we may define

$$e_{\mu}{}^{\nu}(\mathrm{x}): \quad \mathrm{e}_{\mu} = e_{\mu}{}^{\nu}(\mathrm{x}) \frac{\partial}{\partial x^{\nu}}, \tag{B.67}$$

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

exhibiting that the basis elements $\mathrm{e}_\mu(\mathrm{x})$ span a linear vector space. This locally (infinitesimally) defined (tangent) spacetime must then be isomorphic to $\mathbb{R}^{1,3}$, and we are free to choose Cartesian coordinates in it, say ξ^m, for which $g_{mn}(\xi) = -\eta_{mn}$, so that $\mathrm{e}_\mu(\mathrm{x}) = e_\mu{}^m(\mathrm{x})\,\frac{\partial}{\partial\xi^m}$. In turn, comparing the straightforward computation

$$\mathrm{e}_\mu := \frac{\partial \mathrm{e}}{\partial x^\mu} = \frac{\partial \xi^m}{\partial x^\mu}\frac{\partial \mathrm{e}}{\partial \xi^m}, \tag{B.68}$$

with the definition of $e_\mu{}^\nu(\mathrm{x})$ given in (B.67), we see that $e_\mu{}^m(\mathrm{x}) = \frac{\partial\xi^m}{\partial x^\mu}$, when the local tangent-space derivatives $\frac{\partial}{\partial\xi^m}$ are used as covariant basis elements, instead of the curvilinear $\frac{\partial \mathrm{e}}{\partial x^\mu}$.

The so-defined 4×4 matrix of coefficients $e_\mu{}^m(\mathrm{x})$ is variously called a *tetrad*, a *Fierbein* (German: *fier* = four, *Bein* = leg), a “moving frame,” or a “soldering form” [508, 62, 367, 548, 66, 96], as it relates curvilinear derivatives to the local, tangent-space, $\frac{\partial}{\partial x^\mu} = e_\mu{}^m(\mathrm{x})\frac{\partial}{\partial\xi^m}$, at every point in spacetime. Straightforwardly,

$$e_\mu{}^m(\mathrm{x})\,(-\eta_{mn})\,e_\nu{}^n(\mathrm{x}) = g_{\mu\nu}(\mathrm{x}), \tag{B.69}$$

and $e_\mu{}^m(\mathrm{x})$ may be regarded as a square-root of the metric tensor. By abuse of language, one says that $\mu,\nu,\dots$ are “curved indices,” meaning that they indicate curvilinear coordinates; in turn, $m,n,\dots$ are dubbed “flat indices,” meaning that they indicate Cartesian coordinates in the flat tangent spacetime $\cong \mathbb{R}^{1,3}$, which is defined locally (infinitesimally) at every point x of otherwise arbitrarily curved but smooth spacetime.

Clearly, at any point where the local system of partial derivatives $\frac{\partial\xi^m}{\partial x^\mu}$ is ill-defined, this construction in the specified coordinates breaks down, detecting a candidate (putative) singularity; see the discussion in Section 9.3.1, starting on p. 334.

B.2.4 Functionals and functional derivatives

Without delving into technical details and a rigorous definition of functionals and functional derivatives, we provide here a heuristic introduction and a few results that prove useful in computations such as done in Digression 5.9 on p. 191 or Section 11.2.4.

Consider first a 4-vector $\mathrm{x} = (x^0, x^1, x^2, x^3)$. The value of the symbol “k^μ” clearly depends on the choice of the index, which indicates one of the four components. Note that there are only a finite number of choices for μ, and thus a finite number of components of k^μ. This is conceptually similar to the notion of a function $f(x)$, the value of which depends on the choice of the argument x – except that x varies *continuously* over a range of values. For each of the permissible choices of the argument x, $f(x)$ returns a value and so the space of possible values may well also form a continuously infinite set.

We frequently consider summation over the indices – which we will write explicitly in this section, such as,

$$(\mathrm{x}\cdot\eta)_\nu := \sum_{\mu=0}^{3} x^\mu \eta_{\mu\nu} = x^0\eta_{0\nu} + x^1\eta_{1\nu} + x^2\eta_{2\nu} + x^3\eta_{3\nu}. \tag{B.70}$$

In the 4-vector quantity so defined, the index ν appears on both sides of the equation and remains *free*: it may be freely chosen and changed at will. By contrast, the index μ has been summed over, does not even appear in the right-most, expanded version of the sum, and is not free to substitute arbitrarily chosen values (from within $0,1,2,3$); it is a dummy summation variable. *Conceptually,* this is identical to the fact that in the integral

$$F[f;y] := \int_a^b \mathrm{d}x\; f(x)\,H(x,y), \tag{B.71}$$

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

the argument y remains free and available for substitution with any of its allowed values, while the variable x has been "used up" to compute the integral. Just as the sum (B.70) depends on the 4-vector x and its 4 components (x^0, x^1, x^2, x^3), so does the integral (B.71) depend on the choice of the function $f(x)$ and its values. Just as $(\mathrm{x}\cdot\eta)_\nu$ no longer depends on the "used-up" index μ, neither does the integral $F[f;y]$ depend on the "used-up" variable x.

Both of these expressions depend on the summation (integration) limits, they also depend on the additional rank-2 tensor (2-argument function) quantities, $\eta_{\mu\nu}$ $(H(x,y))$, but we focus here the dependence on the 4-vector x^μ vs. the function $f(x)$. In particular, we easily compute the derivative by x^α of the first of these quantities:

$$\begin{aligned}\frac{\partial}{\partial x^\alpha}(\mathrm{x}\cdot\eta)_\nu &= \frac{\partial}{\partial x^\alpha}\sum_{\mu=0}^{3} x^\mu \eta_{\mu\nu} = \sum_{\mu=0}^{3}\frac{\partial}{\partial x^\alpha}x^\mu\eta_{\mu\nu} = \sum_{\mu=0}^{3}\Big(\frac{\partial x^\mu}{\partial x^\alpha}\eta_{\mu\nu} + x^\mu\underbrace{\frac{\partial \eta_{\mu\nu}}{\partial x^\alpha}}_{\text{assume }=0}\Big)\\ &= \sum_{\mu=0}^{3}\Big(\frac{\partial x^\mu}{\partial x^\alpha} = \delta^\mu_\alpha\Big)\eta_{\mu\nu} = \eta_{\alpha\nu}. \end{aligned} \tag{B.72}$$

In the indicated assumption, we state that the rank-2 tensor $\eta_{\mu\nu}$ is defined independently of the 4-vector x^μ. In perfect analogy with (B.72), we compute the *functional* (also called *variational*) derivative of the integral (B.71):

$$\begin{aligned}\frac{\delta}{\delta f(z)}F[f;y] &= \frac{\delta}{\delta f(z)}\int_a^b \mathrm{d}x\, f(x)\,H(x,y) = \int_a^b \mathrm{d}x\,\frac{\delta}{\delta f(z)}f(x)\,H(x,y)\\ &= \int_a^b \mathrm{d}x\,\Big(\frac{\delta f(x)}{\delta f(z)}H(x,y) + f(x)\underbrace{\frac{\delta H(x,y)}{\delta f(z)}}_{\text{assume }=0}\Big)\\ &= \int_a^b \mathrm{d}x\,\Big(\frac{\delta f(x)}{\delta f(z)} = \delta(x-z)\Big)H(x,y) = H(z,y). \end{aligned} \tag{B.73}$$

In the indicated assumption, we state that the 2-argument function $H(x,y)$ is defined independently of the function $f(x)$. Still more generally, consider a nonlinear functional of the function $f(x)$:

$$\mathcal{F}[f] := \int_a^b \mathrm{d}x\,\mathscr{F}\big(f(x)\big), \tag{B.74}$$

where $\mathscr{F}$ is an arbitrary functional expression involving $f(x)$, such as $\frac{(f(x))^2}{\sqrt{\log(f(x)+1)}}$. Requiring the basic chain rule to apply, we obtain

$$\frac{\delta}{\delta f(z)}\mathcal{F}[f] := \int_a^b \mathrm{d}x\,\frac{\delta}{\delta f(z)}\mathscr{F}\big(f(x)\big) = \int_a^b \mathrm{d}x\,\delta(x-z)\Big[\frac{\partial\mathscr{F}(\xi)}{\partial\xi}\Big]_{\xi\to f(x)}, \tag{B.75}$$

where the symbol f is used in the partial derivative within the square brackets as a formal argument of the function $\mathscr{F}$ and the derivative is calculated in the standard way. Once the derivative is computed, $\xi \to f(x)$ is substituted back in the resulting (derivative) functional expression. Note that the Dirac δ-function, $\delta(x-z)$, quenches the integration to an evaluation at $x \to z$.

There is, however, an aspect of functional derivatives that does not have a direct analogue in the 4-vector calculus framework of (B.72), and it has to do with cases where the definition of the functional such as (B.71) depends not only on the function, but also on its derivatives.

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

This is actually a fairly typical case as most Lagrangians are functional expressions involving not only fields, but also their time derivatives. The time-integral of any Lagrangian is then Hamilton's action, and it is a functional of the fields involved. For example,

$$S[\phi;C] = \int_0^T \mathrm{d}t\; L(\phi,\phi',\phi'',\dots). \tag{B.76}$$

$$\begin{aligned}\frac{\delta}{\delta\phi(\tau)}S[\phi;C] &= \int_0^T \mathrm{d}t\; \frac{\delta}{\delta\phi(\tau)}L(\phi,\phi',\phi'',\dots)\\ &= \int_0^T \mathrm{d}t\; \Big\{\frac{\delta\phi(t)}{\delta\phi(\tau)}\Big[\frac{\partial L}{\partial\phi}\Big] + \frac{\delta\phi'(t)}{\delta\phi(\tau)}\Big[\frac{\partial L}{\partial\phi'}\Big] + \frac{\delta\phi''(t)}{\delta\phi(\tau)}\Big[\frac{\partial L}{\partial\phi''}\Big] + \cdots\Big\},\end{aligned} \tag{B.77}$$

where all the expressions in the square brackets treat $\phi, \phi', \phi'', \dots$ as independent variables to perform the indicated partial derivatives, then re-submit $\phi \to \phi(t)$, $\phi' \to \phi'(t)$, etc. Next, we use without proof that $\frac{\delta\phi'(t)}{\delta\phi(\tau)} = \big(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta\phi(t)}{\delta\phi(\tau)}\big)$:

$$\begin{aligned}&= \int_0^T \mathrm{d}t\; \Big\{\frac{\delta\phi(t)}{\delta\phi(\tau)}\Big[\frac{\partial L}{\partial\phi}\Big] + \Big(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\delta\phi(t)}{\delta\phi(\tau)}\Big)\Big[\frac{\partial L}{\partial\phi'}\Big] + \Big(\frac{\mathrm{d}^2}{\mathrm{d}t^2}\frac{\delta\phi(t)}{\delta\phi(\tau)}\Big)\Big[\frac{\partial L}{\partial\phi''}\Big] + \cdots\Big\}\\ &= \int_0^T \mathrm{d}t\; \Big\{\delta(t-\tau)\Big[\frac{\partial L}{\partial\phi}\Big] + \Big(\frac{\mathrm{d}}{\mathrm{d}t}\delta(t-\tau)\Big)\Big[\frac{\partial L}{\partial\phi'}\Big] + \Big(\frac{\mathrm{d}^2}{\mathrm{d}t^2}\delta(t-\tau)\Big)\Big[\frac{\partial L}{\partial\phi''}\Big] + \cdots\Big\}.\end{aligned}$$

Next, we integrate by parts; the second term once, the third term twice and so on:

$$= \int_0^T \mathrm{d}t\; \delta(t-\tau)\Big\{\Big[\frac{\partial L}{\partial\phi}\Big] - \Big(\frac{\mathrm{d}}{\mathrm{d}t}\Big[\frac{\partial L}{\partial\phi'}\Big]\Big) + \Big(\frac{\mathrm{d}^2}{\mathrm{d}t^2}\Big[\frac{\partial L}{\partial\phi''}\Big]\Big) + \cdots\Big\} + \text{B.T.}, \tag{B.78}$$

where "B.T." denotes boundary terms stemming from the integrations by part. Finally,

$$\frac{\delta}{\delta\phi(\tau)}S[\phi;C] = \sum_{k=0}^{\infty}(-1)^k\frac{\mathrm{d}^k}{\mathrm{d}\tau^k}\frac{\partial\, L(\phi(\tau),\phi'(\tau),\phi''(\tau),\dots)}{\partial\,\phi^{(k)}(\tau)} + \text{B.T.} \tag{B.79}$$

A further generalization of this to n-tuples of fields, and to dependence on more than one variable is straightforward:

$$\frac{\delta}{\delta\phi_a(\mathrm{x})}S[\phi_\cdot;C] = \sum_{k=0}^{\infty}(-1)^k\partial^k\frac{\partial\, L(\phi_\cdot(\mathrm{x}),\partial^1\phi_\cdot(\mathrm{x}),\partial^2\phi_\cdot(\mathrm{x}),\dots)}{\partial(\partial^k\phi_a^{(k)}(\mathrm{x}))} + \text{B.T.}, \tag{B.80a}$$

$$\partial^k := \underbrace{\partial_\mu\partial_\nu\cdots\partial_\rho}_{k\ \text{factors}}, \qquad \text{and} \qquad a = 1,2,\dots,n, \tag{B.80b}$$

and where a summation is implied between each spacetime partial derivative occurring within the two copies of ∂^k – one acting on the partial derivative of the Lagrangian density and the other in the specification of the derivative field with respect to which the partial derivative of the Lagrangian is computed:

$$= \frac{\partial L}{\partial\phi_a(\mathrm{x})} - \partial_\mu\frac{\partial L}{\partial(\partial_\mu\phi_a(\mathrm{x}))} - \partial_\mu\partial_\nu\frac{\partial L}{\partial(\partial_\mu\partial_\nu\phi_a(\mathrm{x}))} + \cdots + \text{B.T.} \tag{B.80c}$$

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

B.2.5 Exercises for Section B.2

✎ **B.2.1** *Using the results (9.4) and (3.11c), prove that the line element (B.27) is invariant under general coordinate transformations.*

✎ **B.2.2** *Using the standard rules of calculus (B.23)–(B.24), prove that (B.28) is invariant under general coordinate transformations.*

✎ **B.2.3** *Using the standard rules of calculus (B.23)–(B.24), prove that the Levi-Civita symbol (B.29)–(B.30) is invariant under general coordinate transformations.*

✎ **B.2.4** *Prove the relationship (B.31).*

B.3 A telegraphic introduction to Gödelian incompleteness

This sketchy and perforce incomplete account of Kurt Gödel's incompleteness theorem and its corollary, as well as their implications for all sufficiently complex theoretical systems, is meant to alleviate the fact that most physics students are not familiar with it. For a more complete and precise introduction, see Refs. [211, 376].

With excellent prospects of hilariously oversimplifying the historical background and significance of Gödel's theorem, let me just mention as a backdrop the incredible *Principia Mathematica* by A. N. Whitehead and B. Russell: This three-tome opus [571, 568, 569, $\sim$2,000 pages in total], justifying even a 500-page abridged version of Vol. 1 [570], collects the best efforts to cast the complete and rigorous foundation of all mathematics in Peano's formal symbolic logic and Frege's set theory. The first edition of the *Principia Mathematica* was published in 1910, and was then improved for the second edition in 1927.

The ultimate hope was that all of mathematics could be shown to be deducible from an effectively generable collection of axioms,[4] and by means of perfectly rigorous logic. Whitehead and Russell's opus not only set formidable standards for the rigor of proof (hereafter to be pursued in mathematics), but provided an indelible influence on a century of development in (mathematical) logic and set theory, and *metamathematics* – the mathematics of how mathematics is to be practiced and understood.

In 1931, Kurt Gödel published an announcement of his incompleteness theorem, its corollary (often referred to as the second incompleteness theorem) and an elaborate sketch of proof, deferring the complete proof (to the level of rigor as set by the *Principia Mathematica*). His results were, however, accepted at once and Gödel never did get around to publishing the completely detailed proof [211, 376].

Gödel's incompleteness theorem and its corollary pertain to axiomatic systems that are sufficiently complex to contain the axiomatic system of standard arithmetic. Recall that an axiomatic system is a logical system that has, roughly:

1. a fixed list of symbols,
2. a fixed list of "syntactic/grammatical" rules specifying which strings of symbols represent "well-formed" (meaningful) expressions and statements,
3. a fixed list of adopted logical rules of manipulating and combining statements, and
4. a fixed list of "axioms" (postulates) – statements that are adopted as the "primary statements (truths)" of the given system.

[4] A collection of objects is effectively generable if there exists an algorithm that will enumerate all the objects in the collection without ever enumerating anything else.

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

Every such axiomatic system then has statements that are spelled out with its symbols (1) and which are well-formed (2); those that can be derived using the rules (3) from the axioms (4) are called "theorems."

Within this framework, it was the hope of the research culminating with Whitehead and Russell's *Principia Mathematica* that a suitable system of axioms could be found for all of mathematics, such that every well-formed mathematical statement could either be proven (by deriving it from the axioms) or disproven (by deriving its logical negation instead).

Gödel's **incompleteness theorem** states that no axiomatic system that is sufficiently complex to contain arithmetics can both be complete and not be self-contradictory. That is, to avoid being self-contradictory, every such axiomatic system must contain statements that can be neither proven nor disproven within the axiomatic system as given. Gödel also drew an immediate **corollary** (oft-cited as his **second incompleteness theorem** owing to its importance), which states that no such axiomatic system can prove/demonstrate its own consistency [211].

Even more remarkably, Gödel's proof is constructive! Within any such axiomatic system, Gödel's proof explicitly shows how to construct a very specific statement, which can neither be proven nor disproven within the given axiomatic system. Although Gödel constructed this particular undecidable statement in his proof, and expressly for the purpose of proving the theorem, it does follow that there exist infinitely many such undecidable statements – and some of those, within physics as a formal axiomatic system, are bound to be of interest. Such undecidable statements are then often called Gödelian, although strictly speaking this name should be reserved for the specific statement constructed in Gödel's proof for the given axiomatic system.

Conclusion B.1 *To any axiomatic system (sufficiently complex so as to contain arithmetic), either a Gödelian undecidable statement or its logical negation may be added as a new axiom – and this extension may be repeated recursively forever [211, 376].*

It is worth noticing that the Popperian notion of falsifiability (at least in an admittedy naive understanding [☞ Digression 1.1 on p. 9]) presupposes all statements that one may spell out within some theory (or theoretical system) necessarily to be either falsified or confirmed – so that there *must* exist provable/derivable statements within that theory, which then Nature (experiment) could falsify. In turn, a theory may well be *undecided* about any particular and otherwise perfectly self-consistent statement being tested. Nature then *may* choose one of the options, so effectively decide the statement – and extend the theory.

Example B.1 Within the standard theoretical system of Newtonian classical mechanics, Bertrand's theorem [☞ textbooks of classical mechanics such as [213]] guarantees that stable circular orbits in $(3+1)$-dimensional spacetime are ensured only by two central potentials:

1. the Kepler/Newton potential, $-\frac{\varkappa}{r}$,
2. the radial harmonic potential, $\frac{1}{2}kr^2$.

However, there is nothing within this theoretical system that could decide which one is *the* one that keeps the planets in stable and nearly circular orbits around the Sun. It is the correlation between the orbital linear velocities of the planets and their distance from the Sun – observed in Nature to be $v \propto r^{-1/2}$ [✐ *derive*]– that clearly picks the Kepler/Newton potential over the $v \propto r^{3/2}$ [✐ *derive*] of the harmonic potential.

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

Example B.2 Within the standard theoretical system of Newtonian classical mechanics, there exists no reason to impose Bohr's ad hoc quantization of the angular momentum for the electron orbiting the proton and so forming a hydrogen atom. However, neither does there exist a reason *against* such a quantization: strictly speaking, the assumed continuous variability of the magnitude of angular momentum in various physical systems is merely an implicit assumption, bolstered by no noticed exemption in the macroscopic world; see, however, the discussion in Section 8.3.1 and Footnote *11* on p. 310 in particular.

Therefore, from within the formal theoretical system of classical mechanics, whether or not the angular momentum of an electron orbiting a proton is to be quantized and in what units is in fact an undecidable statement in Gödel's sense. Nature quite clearly resolves the issue: Experiments show that the angular momentum of any physical system can only change in integral multiples of $\hbar$, and so must be either an integral or a half-integral multiple of this unit. The quantum extension of classical mechanics is in this sense precisely a Gödelian extension of the axiomatic theoretical system of classical physics to the axiomatic theoretical system of quantum physics.

Example B.3 As discussed in Section 9.1.1 and Digression 8.1 on p. 295, attempting to fuse Newtonian mechanics and Maxwell's electrodynamics requires one to either modify electrodynamics so as to become Galilean-symmetric, or mechanics so as to become Lorentz-symmetric. Since the Galilean group is the $c \to \infty$ limit of the Lorentz group, the former of these options is achievable only if we take the $c \to \infty$ limit of the Maxwell equations. Both resulting systems are consistent, so that a choice between them is not decidable from within the theory alone. It is indeed Nature's "choice" that light does propagate at a finite speed, which then implies the latter option for the electrodynamics of moving electric charges.

While one may wish for such a "resolution by Nature," as described in the Examples B.1, B.2 and B.3 above, there is in fact no guarantee that *all* "theoretical" dichotomies in our attempts to describe Nature will be similarly resolvable by observation. Indeed, the discovery of the ever-increasing list of ever more various dualities [☞ Section 11.4, to begin with] seems to indicate that this "plurality" of description is an innate characteristic of our understanding Nature.

Finally, the prospect of perpetual Gödelian extensions – in as much as it seems applicable to physics – seems to agree with some of the historical lessons, seen with the benefit of hindsight. Within the theoretical system of classical vector fields, the model described by the Maxwell equations is "well-formed," but undecidable. There is nothing in classical field theory formalism that could prove or disprove the Maxwell equations from any system of axioms, which does not in fact include either the electrodynamics laws that those differential equations represent or the gauge principle as introduced in Chapter 5.

In this sense then, the gauge principle (or the electrodynamics laws represented by the Maxwell equations) is a Gödelian undecidable statement within the theoretical system of classical fields. By including the gauge principle, we obtain the particular theoretical system of classical fields that is called electrodynamics, in which the vector fields $\vec{E}$, $\vec{B}$, $\vec{A}$ and the scalar field Φ acquire

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press

a specific meaning and application. The theoretical system at hand has at once both become more specific and acquired a richer structure (less arbitrariness).

Gödel's incompleteness theorem then implies that the axiomatic system of theoretical physics definitely *can* be extended indefinitely, and in infinitely many ways. Which of those extensions will turn out to be useful towards the intended purpose of theoretical physics, of course, remains an open question – and may well remain so indefinitely☝.

https://doi.org/10.1017/9781009291507.018 Published online by Cambridge University Press