Chapter 9

Symmetries and transformation groups

When contemporaries of Galilei argued against the heliocentric world view by pointing out that we do not feel like rotating with high velocity around the sun he argued that a uniform motion cannot be recognized because the laws of nature that govern our environment are invariant under what we now call a Galilei transformation between inertial systems. Invariance arguments have since played an increasing role in physics both for conceptual and practical reasons.

In the early 20th century the mathematician Emmy Noether discovered that energy conservation, which played a central role in 19th century physics, is just a special case of a more general relation between symmetries and conservation laws. In particular, energy and momentum conservation are equivalent to invariance under translations of time and space, respectively. At about the same time Einstein discovered that gravity curves space-time so that space-time is in general not translation invariant. As a consequence, energy is not conserved in cosmology.

In the present chapter we discuss the symmetries of non-relativistic and relativistic kinematics, which derive from the geometrical symmetries of Euclidean space and Minkowski space, respectively. We decompose transformation groups into discrete and continuous parts and study the infinitesimal form of the latter. We then discuss the transition from classical to quantum mechanics and use rotations to prove the Wigner-Eckhart theorem for matrix elements of tensor operators. After discussing the discrete symmetries parity, time reversal and charge conjugation we conclude with the implications of gauge invariance in the Aharonov–Bohm effect.

9.1 Transformation groups

Newtonian mechanics in Euclidean space is invariant under the transformations

$$g_v(t,\vec{x}) = (t, \vec{x}+\vec{v}t) \qquad \text{Galilei transformation,} \tag{9.1}$$

$$g_{\tau,\vec{\xi}}(t,\vec{x}) = (t+\tau, \vec{x}+\vec{\xi}) \qquad \text{time and space translation,} \tag{9.2}$$

$$g_{\mathcal{O}}(t,\vec{x}) = (t, \mathcal{O}\vec{x}) \qquad \text{rotation or orthogonal transformation,} \tag{9.3}$$

where $\mathcal{O}$ is an orthogonal matrix $\mathcal{O}\cdot\mathcal{O}^T = \mathbb{1}$. In special relativity the structure of the invariance group unifies to translations $x^\mu \to x^\mu + \xi^\mu$ and Lorentz transformations

$$x^\mu \to L^\mu{}_\nu x^\nu, \qquad L^T g L = g \quad \text{with} \quad g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{9.4}$$

which leave $x^2 = x^\mu x_\mu$ invariant. A transformation under which the equations of motion of a classical system are invariant is called a symmetry. Transformations, and in particular symmetry transformations, are often invertible and hence form a group under composition.[1]

Infinitesimal transformations. For continuous groups, whose elements depend continuously on one or more parameters, it is useful to consider infinitesimal transformations. Invariance under *infinitesimal* translations, for example, implies invariance under *all translations*. For the group $O(3) \equiv O(3,\mathbb{R})$ of real orthogonal transformations in 3 dimensions this is, however, not true because $\mathcal{O}\cdot\mathcal{O}^T = \mathbb{1}$ only implies $\det\mathcal{O} = \pm 1$ and transformations with a negative determinant $\det\mathcal{O} = -1$, which change of orientation, can never be reached in a continuous process by composing small transformations with determinant $+1$. The orthogonal group $\mathcal{O}(3)$ hence decomposes into two connected parts and its subgroup $SO(3,\mathbb{R})$ of *special* orthogonal matrices R (*special* means the restriction to $\det R = 1$) is the component that contains the identity. In three dimensions every special orthogonal matrix corresponds to a rotation $R_{\vec{\alpha}}$ about a fixed axis with direction $\vec{\alpha}$ by an angle $|\alpha|$ for some vector in $\vec{\alpha} \in \mathbb{R}^3$. Obviously any such rotation can be obtained by a large number of small rotations so that

$$R_{\vec{\alpha}} = (R_{\frac{1}{n}\vec{\alpha}})^n = \lim_{n\to\infty}(\mathbb{1} + \tfrac{1}{n}\delta R_{\vec{\alpha}})^n = \exp(\delta R_{\vec{\alpha}}) \tag{9.5}$$

where we introduced the infinitesimal rotations

$$\delta R(\vec{\alpha})x_i = \varepsilon_{ijk}\alpha_j x_k = \delta R(\vec{\alpha})_{ik} x_k, \qquad \delta R(\vec{\alpha})_{ik} = \varepsilon_{ijk}\alpha_j. \tag{9.6}$$

Like for the derivative f' of a function f in differential calculus, an infinitesimal transformation δT is the linear term in the expansion $T(\varepsilon\alpha) = \mathbb{1} + \varepsilon\delta T(\alpha) + \mathcal{O}(\varepsilon^2)$ and hence is linear and

[1] Since translations $g_{\vec{\xi}}$ and orthogonal transformations $g_{\mathcal{O}}$ do not commute they generate the Euclidean group $E(3)$ as a *semidirect product*, each of whose elements can be written uniquely as a composition $g_{\mathcal{O}} \circ g_{\vec{\xi}}$. Lorentz transformations and translations in Minkowski space similarly generate the *Poincaré group*.

obeys the Leibniz rule for products and the chain rule for functions,[2]

$$\delta T\,(f\cdot g)=\delta T(f)\cdot g+f\cdot\delta T(g),\qquad \delta T(f(x))=\frac{df(x)}{dx^i}\,\delta T(x^i) \tag{9.7}$$

In accord with (9.6) the infinitesimal form δR of an orthogonal transformation $RR^T=\mathbb{1}$ is given by an antisymmetric matrix since $(\mathbb{1}+\varepsilon\delta R)(\mathbb{1}+\varepsilon\delta R^T)=\mathbb{1}+\varepsilon(\delta R+\delta R^T)+\mathcal{O}(\varepsilon^2)$. Similarly, the infinitesimal form $\delta U=iH$ of a unitary transformation $UU^\dagger=\mathbb{1}$ is antihermitian

$$U=\mathbb{1}+\varepsilon\,iH+\mathcal{O}(\varepsilon^2),\qquad UU^\dagger=\mathbb{1}\quad\Rightarrow H=H^\dagger. \tag{9.8}$$

In turn, $\exp(iH)$ is unitary if H is Hermitian. The advantage of infitesimal transformations is that they just add up for combined transformations,

$$T_1T_2=\mathbb{1}+\varepsilon(\delta T_1+\delta T_2)+\mathcal{O}(\varepsilon^2). \tag{9.9}$$

In particular, an infinitesimal rotation about an arbitrary axis $\vec\alpha$ can be written as a linear combination of infinitesimal rotations about the coordinate axes

$$\delta R(\vec\alpha)=\alpha_j\delta R_j,\qquad (\delta R_j)_{ik}=\delta R(\vec e_j)_{ik}=\varepsilon_{ijk}. \tag{9.10}$$

Since the finite transformations are recovered by exponentiation the Baker–Campbell–Hausdorff formula

$$e^A\,e^B=e^{A+B+\frac12[A,B]+\frac1{12}([A,[A,B]]-[B,[A,B]])+\text{ multiple commutators}} \tag{9.11}$$

shows that a nonabelian group structure of finite transformations corresponds to nonvanishing commutators of the infinitesimal transformations. In the following an infintesimal transformation will not always be indicated by a variation symbol, but it should be clear from the context which transformations are finite and which are infinitesimal.

Discrete transformations As we observed for the orthogonal group, invariance under infinitesimal transformations only implies invariance for the connected part of a transformation group and a number of discrete "large" transformations, which cannot be obtained by combining many small transformations, may have to be investigated seperately. In nonrelativistic mechanics the relavant transformations are time reversal $T:t\to -t$ and parity $P:\vec x\to-\vec x$, which is equivalent to a reflecion $\vec x\to\vec x-2\vec n(\vec x\cdot\vec n)$ at a mirror with normalized orthogonal vector $\vec n$ combined with a rotation $R(\pi\vec n)$ about $\vec n$ by the angle π. In 1956 T.D. Lee and C.N. Yang came up with the idea that an apparent problem with parity selection rules in neutral kaon decay might be due to violation of parity in weak interactions and they suggested a number of experiments for testing the conservation of parity in weak processes. By the end of that year

[2] Linear transformations obeying the Leibniz rule on some associative algebra, like the commutative algebra of functions in classical mechanics or the noncommutative algebras of matrices or operators in quantum mechanics are called *derivations*.

Madame C.S. Wu and collaborators observed the first experimental signs of parity violation in β-decay of polarized ^{60}Co. This experimental result came as a great surprise because parity selection rules had become a standard tool in atomic physics and parity conservation was also well established for strong interactions.

In the relativistic theory there is another discrete transformation, called charge conjugation, which amounts to the exchange of particles and anti-particles. The combination CP of parity and charge conjugation turns out to be even more natural than parity alone, and CP is indeed conserved in many weak processes. But in 1964 it was discovered that CP is also violated in the neutral kaon system.[3] In 1967 Sakharov showed that CP-violation, in addition to thermal non-equilibrium and the existence of baryon number violating processes, is one of the three conditions for the possibility of creating matter in the universe. Invariance under the combination CPT of all three discrete transformations of relativistic kinematics, can be shown to follow from basic axioms of quantum field theory, and indeed no CPT violation has ever been observed.

Active and passive transformations. A transformation like, for example, a translation $\vec{x} \to \vec{x}' = \vec{x} + \vec{\xi}$ can be interpreted in two different ways. On the one hand, we can think of it as a motion where a particle located at the position $\vec{x}$ is moved to the position with coordinates $\vec{x}'$ with respect to some fixed frame of reference. Such a motion is often called an *active transformation*. On the other hand we can leave everything in place and describe the same physical process in terms of new coordinates x'. The resulting coordinate transformation is often called a *passive transformation*. Active and passive transformation are mathematically equivalent in the sense that the formulas look identical. If we physically move our experiment to a new lab, however, our instruments may be sufficiently sensitive to detect the change of the magnetic field of the earth or of other environmental parameters that are not moved in an active transformation of the experiment. If we also move the earth and its magnetic field, then it is most likely that we are still in our old lab and that all that happened was a change of coordinates.

If we simultaneously perform an active and a passive transformation then a scalar quantity like a wave function $\psi(x)$ does not change its form so that

$$\psi(x) = \psi'(x'), \qquad x' = Rx \quad \Rightarrow \quad \psi'(x) = \psi(R^{-1}x). \tag{9.12}$$

Quantum mechanics has its own way of incorporating this relation into its formalism. Since a symmetry transformation has to preserve scalar products we consider unitary transformations

[3] Since quarks have both weak and strong interactions, it is still mysterious why CP violation does not also affect the strong interactions. This is known as the *strong CP problem*. Its only proposed explanation so far has been the Peccei-Quinn symmetry, which postulates a new particle called axion. If they exist, axions might contribute to the observed dark matter in the universe.

$R^\dagger = R^{-1}$ in Hilbert space. Hence

$$(R\psi)(x) = \langle x|R|\psi\rangle = \langle R^\dagger x|\psi\rangle = \langle R^{-1}x|\psi\rangle = \psi(R^{-1}x), \tag{9.13}$$

in accord with (9.12). For discrete symmetries also anti-unitary maps are possible, as will be the case for time reversal and charge conjugation.

9.2 Noether theorem and quantization

Canonical mechanics. For a dynamical system with *Lagrange function* $L = L(q^i, \dot q^i, t)$ *Hamilton's principle of least action* states that the functional

$$\phi(\gamma) = \int_{t_0}^{t_1} \mathrm{d}t L(q^i, \dot q^i, t) \tag{9.14}$$

has to be extremal among all paths $\gamma = \{q(t)\}$ with fixed initial point $q^i(t_0)$ and fixed final point $q^i(t_1)$. Since $\delta\dot q^i = \frac{d}{dt}\delta q^i$ the variation can be written as

$$\delta\phi = \int_{t_0}^{t_1} \mathrm{d}t \left(\frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial \dot q^i}\delta \dot q^i \right) = \int_{t_0}^{t_1} \mathrm{d}t \left((\frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot q^i})\delta q^i + \frac{d}{dt}(\frac{\partial L}{\partial \dot q^i}\delta q^i) \right). \tag{9.15}$$

Due to the boundary conditions the variation $\delta q^i(t)$ is zero at the initial and at the final time so that the surface term $\int dt \frac{d}{dt}(\frac{\partial L}{\partial \dot q^i}\delta q^i) = (\frac{\partial L}{\partial \dot q^i}\delta q^i)|_{t_0}^{t_1}$ vanishes. Extremality of the action $\delta\phi = 0$ for all variations is hence equivalent to the *Euler-Lagrange equations of motion*

$$\frac{\delta L}{\delta q^i} = 0 \qquad \text{with} \qquad \frac{\delta L}{\delta q^i} \equiv \frac{\partial L}{\partial q^i} - \frac{d}{dt}\frac{\partial L}{\partial \dot q^i} = \frac{\partial L}{\partial q^i} - \dot p_i \tag{9.16}$$

where we introduced the *variational derivative* $\frac{\delta L}{\delta q^i}$ of L and the *canonical momentum* $p_i \equiv \frac{\partial L}{\partial \dot q^i}$. The space parametrized by the *canonical coordinates* q^i is called *configuration space*.

By *Legendre transformation* with respect to $\dot q^i$ we obtain the *Hamilton function*

$$H(p_i, q^i, t) = \sum p_i \dot q^i - L(q^i, \dot q^i, t) \qquad \text{with} \qquad p_i = \frac{\partial L}{\partial \dot q^i}, \tag{9.17}$$

as a function of the momenta p_i and the coordinates q^i, which together parametrize the *phase space*. Since the inverse Legendre transformation is given by eliminating the momenta p_i from the equation $\dot q^i = \partial H/\partial p_i(p,q)$ the *Hamiltonian equations of motion*

$$\dot q^i = \frac{\partial H}{\partial p_i}, \qquad \dot p_i = -\frac{\partial H}{\partial q^i} \tag{9.18}$$

are equivalent to the Euler-Lagrange equations. The equations (9.18) can also be obtained directly as variational equations $\delta\tilde L/\delta p_i = 0$ and $\delta\tilde L/\delta q^i = 0$ of the *first order Lagrangian* $\tilde L(q, \dot q, p) = \dot q^i p_i - H(p,q)$. Infinitesimal time evolution, as any infinitesimal transformation, obeys the Leibniz rule for products and the chain rule for phase space functions $f(q,p,t)$, for which we admit an explicit time dependence. Regarding f as a function of time on a classical trajectory we hence obtain

$$\dot f = \frac{\partial f}{\partial q^i}\dot q^i + \frac{\partial f}{\partial p_i}\dot p_i + \partial_t f = \{H, f\}_{PB} + \partial_t f \tag{9.19}$$

where we defined the *Poisson brackets*

$$\{f,g\}_{PB} \equiv \sum_i \left(\frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i}\right) \tag{9.20}$$

for arbitrary phase space functions $f(p,q)$ and $g(p,q)$.

The Noether theorem. An infinitesimal transformation $q^i \to q^i + \hat\delta q^i$ with $\hat\delta q^j = f^j(q^i, \dot q^i)$ is a symmetry of a dynamical system with Lagrange function $L(q^i, \dot q^i)$ if $\hat\delta L = \dot K$ is a total time derivative because a total derivative does not contribute to the variation (9.15) and hence leaves the equations of motion invariant. The Noether theorem states that these infinitesimal symmetries are in one-to-one correspondence with *constants of motion* Q, which are also called *conserved charges* or *first integrals*. More explicitly, a symmetry $\hat\delta q$ with $\hat\delta L = \dot K$ implies that

$$Q = \hat\delta q^i p_i - K \tag{9.21}$$

is a constant of motion. In turn, if some phase space function $Q(q^i, \dot q^i)$ is a constant of motion for all classical trajectories then its time derivative is a linear combination of the equations of motion $\dot Q = \sum_i \rho^i \frac{\delta L}{\delta q^i}$. The transformation $\hat\delta q^i = -\rho^i(q^j, \dot q^j)$ is then a symmetry of the Lagrange function, i.e. $\hat\delta L = \dot K$ with $K = \hat\delta q^i p_i - Q$.

Remarks: A constant of motion is only constant for motions that obey the equations of motion! It is important to discern *identities* and *dynamical equations*. For a constant of motion $\dot Q = 0$ is a consequence of the equations of motion. This implies that there is an *identity* $\dot Q = \sum_i \rho^i \frac{\delta L}{\delta q^i}$ that holds for *arbitrary* functions $q^i(t)$ and not only for solutions to $\frac{\delta L}{\delta q^i} = 0$. A symmetry transformation $\hat\delta q^i$ (like e.g. a translation) does not have to vanish at an initial or final time!

Proof: According to (9.15) the equation

$$\delta L \;=\; \frac{\partial L}{\partial q^i}\delta q^i + \frac{\partial L}{\partial \dot q^i}\delta \dot q^i \;\equiv\; \frac{\delta L}{\delta q^i}\delta q^i + \frac{d}{dt}\Big(\frac{\partial L}{\partial \dot q^i}\delta q^i\Big) \tag{9.22}$$

is an *identity* for an arbitrary variation. $\hat\delta L = \dot K$ hence implies

$$\frac{\delta L}{\delta q^i}\hat\delta q^i + \frac{d}{dt}(p_i \hat\delta q^i) = \dot K. \tag{9.23}$$

The theorem follows by subtracting the time derivative $\frac{d}{dt}(p_i\hat\delta q^i)$ from this equation. □

Hamiltonian version. Coordinate transformations in phase space $(p,q) \to (p', q')$ with functions $p'(p,q)$ and $q'(p,q)$ that leave the form of the Poisson brackets invariant are called *canonical transformations.* It can be shown that infinitesimal canonical transformations $\hat\delta$ can be written in terms of a generating function $Q(p,q)$ as

$$\hat\delta q^i = \{Q,q^i\}_{PB}, \qquad \hat\delta p_i = \{Q,p_i\}_{PB}. \tag{9.24}$$

For a fixed phase space function Q the map $g \to \{Q,g\}_{PB}$ is linear and obeys the Leibniz rule, as required for an infinitesimal transformation. In the canonical formalism a symmetry transformation is, by definition, a canonical transformation $\{Q,.\}_{PB}$ with some generating function $Q(p,q)$ that leaves the Hamilton function invariant $\{Q,H\}_{PB} = 0$. This makes the Noether theorem quite trivial, because $\{Q,H\}_{PB} = -\{H,Q\}_{PB} = -\dot Q = 0$ is at the same time the condition for Q to be a constant of motion.

The equivalence of the Hamiltonian and the Lagrangian definition of a symmetry, as well as the equality of the Noether charges Q, can be seen by computing a variation $\hat\delta$ of the first order Lagrangian $\tilde L = \dot q^i p_i - H$,

$$\begin{aligned}\hat\delta(\dot q^i p_i - H) &= \hat\delta\dot q^i\, p_i + \dot q^i \hat\delta p_i - \hat\delta H &(9.25)\\ &= \frac{d}{dt}(\hat\delta q^i\, p_i) - \hat\delta q^i \dot p_i + \dot q^i\hat\delta p_i - \partial_{q^i} H \hat\delta q^i - \partial_{p_j} H\hat\delta p_j, &(9.26)\end{aligned}$$

which can only be equal to $\dot K(p,q)$ if $\frac{d}{dt}(K - p_i\hat\delta q^i) = \dot q^i\partial_{q^i}(K - p_i\hat\delta q^i) + \dot p_i\partial_{p_i}(K - p_i\hat\delta q^i)$ is equal to $-\hat\delta q^i\dot p_i + \dot q^i\hat\delta p_i$ so that $\hat\delta$ is a canonical transformation

$$\hat\delta q^i = \frac{\partial Q}{\partial p_i} \quad \text{and} \quad \hat\delta p_i = -\frac{\partial Q}{\partial q^i} \qquad \Rightarrow \qquad \hat\delta f = \{Q,f\}_{PB} \tag{9.27}$$

with generating function $Q = p_i\hat\delta q^i - K$. The last two terms in (9.26), which do not contain time-derivatives and hence have to cancel each other, now combine to $\hat\delta H = \{Q,H\} = 0$. We thus have shown that the Noether charge Q, when expressed as a function of the canonical coordinates q^i and p_i, is the generating function for the transformation $\hat\delta$.

Quantization. Since $\{p_i,x_j\}_{PB} = \delta_{ij}$ in classical mechanics and $[P_i,X_j] = \frac{\hbar}{i}\delta_{ij}$ in quantum mechanics canonical quantization replaces Poisson brackets of conjugate phase space variables by $\frac{i}{\hbar}$ times the commutator of the corresponding operators,

$$\{p_i,x_j\}_{PB} = \delta_{ij} = \frac{i}{\hbar}[P_i,X_j]. \tag{9.28}$$

For the generating functions of infinitesimal transformations this amounts to

$$\hat\delta q^i = \{Q,q^i\}_{PB} \qquad \to \qquad \hat\delta\vec X = \frac{i}{\hbar}[Q,\vec X]. \tag{9.29}$$

Note that Poisson brackets and commutators both are antisymmetric, satisfy the Jacobi-Identity, and obey the Leibniz rule for each of its two arguments, so that $\{Q,.\}_{PB}$ and $[Q,.]$

both qualify as inifinitesimal transformations. Moreover, the *real* variation $\hat{\delta}q^i$ of the real coordinate q^i naturally leads to the anti-Hermitian operator $\frac{i}{\hbar}Q$ so that the finite transformation $\exp(\frac{i}{\hbar}Q)$ becomes a unitary operator. More precisely, one has to be careful about possible ordering ambiguities if $Q(q^i, p_i)$ is a composite operator. It is always possible to choose Q Hermitian (for example by $Q \to \frac{1}{2}(Q + Q^\dagger)$, or by Weyl ordering). In quantum mechanics it is usually also possible of find a proper quantum version of the symmetry generators, but for an infinite numbers of degrees of freedom quantum violations of classical symmetries, which are called anomalies, can be unavoidable and may lead to important restrictions for the structure of consistent theories.[4]

Energy, momentum and angular momentum. As an example we now compute the generators of translations and rotations. Under a time translation $\hat{\delta}q^i = \dot{q}^i$ of an autonomous system the Lagrange function transforms into its time derivative $\hat{\delta}L = \dot{K} = \dot{L}$, so that the corresponding Noether charge $Q = \hat{\delta}q^i p_i - K = \dot{q}^i p_i - L$ agrees with the Hamilton function. This proves the equivalence of time independence and energy conservation. Uppon quantization (9.29) we find

$$\frac{d}{dt}\vec{X} = \frac{i}{\hbar}[H, \vec{X}], \tag{9.30}$$

which is Heisenberg's equation of motion for the position operator $\vec{X}$. Canonical quantization hence naturally leads to the Heisenberg picture. The corresponding time evolution of the wave function in the Schrödinger picture is given by the Schrödinger equation $\frac{d}{dt}|\psi\rangle = -\frac{\hbar}{i}H|\psi\rangle$.

The generator of a translation $\hat{\delta}_i\vec{x} = \vec{e}_i$ into the coordinate direction $\vec{e}_i$ is the momentum p_i because $\hat{\delta}_i L = 0$, hence $K = 0$, and $\hat{\delta}_i x_j p_j = \vec{e}_i\vec{p} = p_i$ for a translation invariant Lagrange function. Under rotations a centrally symmetric action is also strictly invariant $\hat{\delta}_{\vec{\alpha}}L = \dot{K} = 0$. A rotation about the x_j-axis is given by $\hat{\delta}\vec{x} = \delta R_{e_j}\vec{x}$. With (9.10) we have $\hat{\delta}x_i = \varepsilon_{ijk}x_k$ and thus obtain the corresponding Noether charge

$$L_j = \hat{\delta}x_i p_i - 0 = \varepsilon_{ijk}x_k p_i = \varepsilon_{jki}x_k p_i \tag{9.31}$$

or $\vec{L} = \vec{x}\times\vec{p}$ in accord with the usual definition of angular momentum. The results are collected in the following table.

Symmetry	Noether charge	infinitesimal transformation
time evolution	Hamiltonian H	$\frac{d}{dt}\|\psi\rangle = -\frac{i}{\hbar}H\|\psi\rangle$
translation	momentum P_i	$-\vec{\nabla}\|\psi\rangle = -\frac{i}{\hbar}\vec{P}\|\psi\rangle$
rotation	orbital angular momentum L_i	$\delta R_\alpha\|\psi\rangle = -\frac{i}{\hbar}\vec{\alpha}\vec{L}\|\psi\rangle$

[4] In the standard model of particle interaction, for example, cancellation of certain anomalies between quarks and leptons is indespensible for the consisteny of the theory, while the anomalay in baryon number conservation is in principle observable and enables proton decay, one of Sakharov's conditions for the creation of matter. Anomalies are also the origin of the space-time dimension 10 in superstring theory.

For finite transformations $U = \exp(-\frac{i}{\hbar}H)$ is the time evolution operator and $\exp(-\frac{i}{\hbar}\vec{a}\vec{P})\psi(x)$ yields the Taylor series expansion of $\psi(\vec{x} - \vec{a})$ in the translation vector $\vec{a}$. For infinitesimal rotations

$$\delta R_{\vec{\alpha}}\psi(x) \equiv -\frac{i}{\hbar}\vec{\alpha}\vec{L}\psi(x) = -\frac{i}{\hbar}\alpha^i\varepsilon_{ijk}x_j\frac{\hbar}{i}\nabla_k\psi(x) = -\varepsilon_{ijk}\alpha_j x^k\nabla_i\psi(x), \tag{9.32}$$

in accord with (9.13).

Transformation of operators. A unitary transformation $|\psi\rangle \to T|\psi\rangle$ of states implies that the respective transformation of operators $\mathcal{O}$ is

$$|\psi\rangle \to T|\psi\rangle \qquad \Rightarrow \qquad \mathcal{O} \to T\mathcal{O}T^\dagger \tag{9.33}$$

because matrix elements should not change if we apply both transformations simultaneously. Since $T\mathcal{O}T^{-1} = (\mathbb{1} + \varepsilon\delta T)\mathcal{O}(\mathbb{1} - \varepsilon\delta T) + \mathcal{O}(\varepsilon^2) = \mathbb{1} + \varepsilon[\delta T, \mathcal{O}] + \mathcal{O}(\varepsilon^2)$ the infinitesimal version of this correspondence is

$$|\psi\rangle \to \delta T|\psi\rangle \qquad \Rightarrow \qquad \delta\mathcal{O} = [\delta T, \mathcal{O}]. \tag{9.34}$$

By the active–passive equivalence, an operator transformation $\mathcal{O} \to T\mathcal{O}T^\dagger$ can hence be replaced by the inverse transformation of states, projectors and density matrices

$$|\psi\rangle \to T^\dagger|\psi\rangle \qquad \Rightarrow \qquad P_\psi \to T^\dagger P_\psi T \quad \text{and} \quad \rho \to T^\dagger\rho T \qquad \text{with} \quad P_\psi = |\psi\rangle\langle\psi|, \tag{9.35}$$

which transforms expectation values $\operatorname{tr} P_\psi\mathcal{O} \to \operatorname{tr}(T^\dagger P_\psi T)\mathcal{O} = \operatorname{tr} P_\psi(T\mathcal{O}T^\dagger)$ in the same way. These rules hold for all unitary transformations and not only for symmetry transformations!

9.3 Rotation of spins

If we consider the total angular momentum operator $\vec{J} = \vec{L} + \vec{S}$ for a particle with spin $\vec{S}$ we obtain its finite rotations by

$$e^{-\frac{i}{\hbar}\vec{\alpha}\vec{J}} = e^{-\frac{i}{\hbar}\vec{\alpha}\vec{L}}e^{-\frac{i}{\hbar}\vec{\alpha}\vec{S}} \tag{9.36}$$

because the orbital angular momentum and the spin operator commute. The former operator is responsible for the shift in the position resulting from the rotation while the latter rotates the orientation of the spin. The operator $\exp(-\frac{i}{\hbar}\vec{\alpha}\vec{S})$ is hence called the rotation operator in *spin space*, and it is often sufficient to study its action if the spin orientation rather than the precise position of the particle is relevant for a computation. In the basis $|s,\mu\rangle$ where S^2 and S_z are diagonal we have

$$S_\pm|s,\mu\rangle = \hbar\sqrt{(s \mp \mu)(s \pm \mu + 1)}\;|s,\mu \pm 1\rangle \qquad \text{and} \qquad S_z|s,\mu\rangle = \hbar\mu\;|s,\mu\rangle \tag{9.37}$$

with $S_\pm = S_x \pm iS_y$ or $S_x = \frac{1}{2}(S_+ + S_-)$ and $S_y = \frac{1}{2i}(S_+ - S_-)$.

Spinors. For spin $s = \frac{1}{2}$ the wave function in the S_z-basis (9.37) can be written as

$$\psi(x) \equiv \psi_+(x)\,|\uparrow\rangle + \psi_-(x)\,|\downarrow\rangle \equiv \begin{pmatrix}\psi_+(x)\\ \psi_-(x)\end{pmatrix} \tag{9.38}$$

with $|\uparrow\rangle = |\frac{1}{2},\frac{1}{2}\rangle = \begin{pmatrix}1\\0\end{pmatrix}$ and $|\downarrow\rangle = |\frac{1}{2},-\frac{1}{2}\rangle = \begin{pmatrix}0\\1\end{pmatrix}$ and the spin operator becomes $\vec S = \frac{\hbar}{2}\vec\sigma$ in terms of the Pauli matrices

$$\sigma_x = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \quad \sigma_y = \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}, \quad \sigma_z = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \qquad \text{with} \qquad \begin{array}{l}\sigma_i\sigma_j = \delta_{ij} + i\varepsilon_{ijk}\sigma_k\\ \mathrm{tr}\,\sigma_i = 0\end{array}. \tag{9.39}$$

Since $(\vec\sigma\vec\alpha)^2 = \alpha_i\alpha_j\sigma_i\sigma_j = \alpha_i\alpha_j\delta_{ij}\mathbb{1} = \alpha^2\mathbb{1}$ exponentiation of $\delta R_{\vec\alpha}|\psi\rangle = -\frac{i}{\hbar}\vec\alpha\vec S|\psi\rangle = -\frac{i}{2}\vec\alpha\vec\sigma|\psi\rangle$ yields the finite spin rotations

$$R_\alpha = e^{-\frac{i}{2}\vec\alpha\vec\sigma} = \mathbb{1}\cos\frac{\alpha}{2} - i\,\vec e_\alpha\vec\sigma\,\sin\frac{\alpha}{2} \qquad \text{with} \qquad \alpha = |\vec\alpha| \quad \text{and} \quad \vec e_\alpha = \vec\alpha/\alpha, \tag{9.40}$$

which leave the position invariant but mix the spin-up and the spin-down components of the wave function. We observe that a rotation by an angle $\alpha = 2\pi$ transforms $|\psi\rangle \to -|\psi\rangle$. This strange behavior of spinors is not inconsistent because $\pm|\psi\rangle$ only differ by a phase and hence represent the same physical state of the system and cannot be distinguished by any observable. Phases do become observable, however, in interference pattens. The change of sign for a rotation by 2π has indeed been verified experimentally with neutrons interferometry by H. Rauch et al. in 1975, who achieved destructive interference between coherent neutron rays whose spins were rotated by a relative angle 2π.

Remark: The projector $\Pi_{|\uparrow,\vec n\rangle}$ onto a state with spin up in the direction of a unit vector $\vec n$ can be obtained by an active rotation $R_{\vec\alpha}$ of $|\uparrow\rangle\langle\uparrow| = \frac{1}{2}(\mathbb{1} + \sigma_z)$ with $\vec\alpha = \frac{\alpha}{\sin\alpha}\vec e_z \times \vec n$ for $\cos\alpha = n_z$,

$$R_{\vec\alpha} = \frac{(1+n_z)\mathbb{1} + in_y\sigma_x - in_x\sigma_y}{\sqrt{2(1+n_z)}}, \qquad \Pi_{|\uparrow,\vec n\rangle} \equiv |\uparrow,\vec n\rangle\langle\uparrow,\vec n| = R_{\vec\alpha}\frac{\mathbb{1}+\sigma_z}{2}R^\dagger_{\vec\alpha} = \frac{1}{2}(\mathbb{1} + \vec n\vec\sigma). \tag{9.41}$$

Without lengthy calculation the result directly follows from the fact that $\vec n\vec\sigma = \frac{2}{\hbar}\vec n\vec S$ has eigenvalues ± 1 on states with spin component $\pm\frac{\hbar}{2}$ in the direction $\vec n$.

Vectors. For spin $s = 1$ the analog of the Pauli matrices can again be obtained from (9.37) with $S_x = \frac{1}{2}(S_+ + S_-)$ and $S_y = \frac{1}{2i}(S_+ - S_-)$,

$$S_x^{(1)} = \frac{\hbar}{\sqrt 2}\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 1\\ 0 & 1 & 0\end{pmatrix} \qquad S_y^{(1)} = \frac{\hbar}{i\sqrt 2}\begin{pmatrix}0 & 1 & 0\\ -1 & 0 & 1\\ 0 & -1 & 0\end{pmatrix} \qquad S_z^{(1)} = \hbar\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & -1\end{pmatrix}. \tag{9.42}$$

In order to relate the spherical basis $|1,m\rangle$ to the standard vector basis $\vec e_i$ we start with

$$\vec e_z = |1,0\rangle \tag{9.43}$$

because $\vec{e}_z$ is an eigenvector of the infinitesimal rotation $\delta R_{\vec{e}_z}$ about the z-axis with eigenvalue 0. Evaluating $S_\pm = S_x \pm iS_y$ on $|1,0\rangle = \vec{e}_z$ we find

$$S_\pm|1,0\rangle = \hbar\sqrt{(1\mp 0)(1\pm 0+1)}|1,\pm 1\rangle = \sqrt{2}\hbar|1,\pm 1\rangle \stackrel{!}{=} (S_x \pm iS_y)\vec{e}_z = -\frac{\hbar}{i}(-\vec{e}_y \pm i\vec{e}_x) \quad (9.44)$$

because $-\frac{i}{\hbar}S_i$ generates an infinitesimal rotation about $\vec{e}_i$. Equality of these expressions implies

$$|1,\pm 1\rangle = \mp\frac{1}{\sqrt{2}}(\vec{e}_x \pm i\vec{e}_y). \quad (9.45)$$

The *spherical components* $V_q^{(1)}$ of a vector $\vec{V}$ in the S_z-basis $|1,m\rangle$ are hence defined by

$$V_0^{(1)} = V_z, \qquad V_{\pm 1}^{(1)} = \mp\frac{1}{\sqrt{2}}(V_x \pm iV_y). \quad (9.46)$$

Since $V_1^{(1)}W_{-1}^{(1)} + V_{-1}^{(1)}W_1^{(1)} = -V_xW_x - V_yW_y$ the scalar product of two vectors becomes

$$\vec{V}\cdot\vec{W} = \sum_{q=-1}^{1}(-1)^q\, V_q^{(1)}W_{-q}^{(1)} \quad (9.47)$$

in the spherical basis. As the matrices $S_i^{(1)}$ in (9.42) neither anticommute nor square to $\mathbb{1}$ there is no simple analog of the formula (9.40) for finite rotations $\exp(-\frac{i}{\hbar}\vec{\alpha}\vec{S}^{(1)})$. A general formula for arbitrary spin is known, however, if we represent the rotation in terms of the Euler angles.

General representation of the rotation group. Every rotation in $\mathbb{R}^3$ can be written as a combination of a rotation by an angle γ about the z-axis followed by a rotation by β about the y-axis and a rotation by α about the z-axis. The angles (α,β,γ) are called *Euler* angles (see appendix A.10 in [Grau]) and the corresponding rotation operator is

$$R = e^{-\frac{i}{\hbar}J_z\alpha}\cdot e^{-\frac{i}{\hbar}J_y\beta}\cdot e^{-\frac{i}{\hbar}J_z\gamma}. \quad (9.48)$$

Since $J_z|j,m\rangle = \hbar m|j,m\rangle$ the matrix elements of this operator in an eigenbasis of J^2 and J_z can be written as

$$\langle j,m'|R|j,m\rangle = e^{-im'\alpha}\langle j,m'|e^{-\frac{i}{\hbar}J_y\beta}|j,m\rangle\, e^{-im\gamma} \quad (9.49)$$

For the non-diagonal rotation operator about the y axis we define

$$d_{m'm}^{(j)}(\beta) = \langle j,m'|e^{-\frac{i}{\hbar}J_y\beta}|j,m\rangle \quad (9.50)$$

Without proof we state the formula

$$d_{m'm}^{(j)}(\beta) = (-1)^{m'-m}\sqrt{\frac{(j-m')!(j+m')!}{(j-m)!(j+m)!}}\sin^{m'-m}\left(\frac{\beta}{2}\right)\cos^{m'+m}\left(\frac{\beta}{2}\right)P_{j-m'}^{m'-m,m'+m}(\cos\beta) \quad (9.51)$$

where $P_n^{r,s}(\xi)$ are the *Jacobi* polynomials, which can be defined by [Grau, A2]

$$P_n^{r,s}(\xi) = \tfrac{(n+r)!}{(n-r)!}\left(\tfrac{1+\xi}{2}\right)^2 F\left(-n,-n-s,r+1;\tfrac{\xi+1}{\xi-1}\right) \quad (9.52)$$

$$= \frac{(-1)^n}{2^n\,n!}(1-\xi)^{-r}(1+\xi)^{-s}\frac{d^n}{d\xi^n}\Big((1-\xi)^{n+r}(1+\xi)^{n+s}\Big) \quad (9.53)$$

in terms of the hypergeometric function F.

SO(3) and SU(2). The tensor product $\langle\varphi|\otimes|\psi\rangle$ of two spinors corresponds to a 2×2 matrix with 4 degrees of freedom, which we expect to contain a scalar and vector. Since the three traceless Pauli matrices and the unit matrix together form a basis for all 2×2 matrices $\langle\varphi|\otimes|\psi\rangle$ can be written as linear combinations of

$$\langle\varphi|\mathbb{1}|\psi\rangle \qquad \text{and} \qquad \langle\varphi|\vec\sigma|\psi\rangle. \tag{9.54}$$

These matrix elements indeed transform as a skalar and a vector, respectively, since

$$\delta R_\alpha(\langle\varphi|\sigma_i|\psi\rangle) = -\frac{i}{2}\alpha_j(\langle\varphi|\sigma_i\sigma_j - \sigma_j\sigma_i|\psi\rangle) = -\frac{i}{2}\alpha_j\langle\varphi|2i\varepsilon_{ijk}\sigma_k|\psi\rangle = \alpha_j\varepsilon_{ijk}\langle\varphi|\sigma_k|\psi\rangle \tag{9.55}$$

and $\delta R_\alpha(\langle\varphi|\psi\rangle) = \frac{i}{2}(\langle\varphi|\alpha\sigma)|\psi\rangle - \frac{i}{2}\langle\varphi|(\alpha\sigma|\psi\rangle) = 0$. Since $\vec\alpha\vec\sigma$ is an arbitrary traceless Hermitian matrix the exponential $A = \exp(-\frac{i}{2}\vec\alpha\vec\sigma)$ is a arbitrary special unitary matrix $A\in SU(2)$ which can be written as

$$SU(2)\ni A = \begin{pmatrix} a & b\\ c & d\end{pmatrix}, \qquad AA^\dagger = A\begin{pmatrix} a^* & c^*\\ b^* & d^*\end{pmatrix} = \mathbb{1} \quad\Rightarrow\quad \begin{matrix}|a|^2+|b|^2=1\\ ca^*+db^*=0\end{matrix}. \tag{9.56}$$

The last equation implies $c = -\frac{b^*}{a^*}d$ so that $\det A = 1 = ad - bc = \frac{d}{a^*}(aa^*+bb^*) = \frac{d}{a^*}$. We thus obtain $d = a^*$, $c = -b^*$ and

$$A = \begin{pmatrix} a & b\\ -b^* & a^*\end{pmatrix} \qquad \text{with} \qquad |a|^2+|b|^2 = 1. \tag{9.57}$$

As a manifold $SU(2)$ is therefore a 3-sphere with radius 1 in $\mathbb{R}^4$ whose coordinates are the real and imaginary parts of a and b. For finite transformations the spinor rotation $|\psi\rangle\to A|\psi\rangle$ leads to the vector rotation

$$\langle\varphi|\sigma_i|\psi\rangle \quad\to\quad \langle\varphi|A^\dagger\sigma_i A|\psi\rangle = \sum_k R_{ik}(A)\langle\varphi|\sigma_k|\psi\rangle. \tag{9.58}$$

The equation

$$A^\dagger\sigma_i A = R_{ik}(A)\sigma_k \tag{9.59}$$

hence defines a map from $A\in SU(2)$ to a rotation $R(A)\in SO(3)$. This map is two-to-one because A and $-A$ lead to the same rotation. This should not come as a surprise because we already know that a rotation by 2π, which is the identity in $SO(3)$, reverses the sign of a spinor. A and $-A$ are antipodal points of the S^3 that represents $SU(2)$. We can therefore think of $SO(3)$ as the 3-sphere with antipodal points identified and $SU(2)$ is a smooth double cover of the rotation group. The mathematical reason for the existence of spinor representations of the rotation group is the fact that $SO(3)$ admits an unbranched double-cover and hence admits so called *projective* or *ray* representations where a full rotation gives back the original state only up to a phase factor. Such objects would be forbidden in classical mechanics, but in quantum mechanics a physical state is not represented by a unique vector $|\psi\rangle\in\mathcal{H}$ but rather by a "ray" of vectors $\lambda|\psi\rangle$ with $\lambda\neq 0$.

9.3.1 Tensor operators and the Wigner Eckhart theorem

Vector and tensor operators are collections of operators labelled by vector or tensor indices that transform accordingly under rotations,

$$[J_i, V_j] = i\hbar\varepsilon_{ijl}V_l, \qquad [J_i, T_{jl}] = i\hbar\varepsilon_{ijm}T_{ml} + i\hbar\varepsilon_{iln}T_{jn}, \qquad \ldots \tag{9.60}$$

The number of indices is the *order* k of the tensor. Since

$$\vec{J} \cdot (T|\psi\rangle) = [\vec{J}, T] \cdot |\psi\rangle + T \cdot \vec{J}|\psi\rangle \tag{9.61}$$

the action of a general vector operator is like an addition of an angular momentum $j = 1$ as far as the properties under rotations are concerned. For a vector operator (9.46) we already know that

$$\left[J_z, V_q^{(1)}\right] = q\hbar V_q^{(1)} \tag{9.62}$$

$$\left[J_\pm, V_q^{(1)}\right] = \sqrt{2 - q(q \pm 1)}\hbar V_{q\pm1}^{(1)} \tag{9.63}$$

For higher order $k > 1$ a general tensor decomposes into irreducible parts that are connected by the ladder operators. An irreducible tensor operator T_q^k of the order k is defined by the following commutation relations,

$$\left[J_z, T_q^{(k)}\right] = \hbar q T_q^{(k)}, \tag{9.64}$$

$$\left[J_\pm, T_q^{(k)}\right] = \sqrt{k(k+1) - q(q \pm 1)}\hbar T_{q\pm1}^{(k)}, \tag{9.65}$$

where k and q are the analogues of the eigenvalues l and m of the spherical harmonics and $q = -k, \ldots, k$ labels the $2k + 1$ spherical components of the irreducible tensor operator $T^{(k)}$.

A tensor operator T_{il} of order 2, for example, has nine elements. Since two spins $j_1 = j_2 = 1$ add up to spin $j \leq 2$ we expect T_{il} to contain irreducible tensors of order $0, 1$ and 2. In this example it is easy to guess that the scalar is the trace $T^{(0)} = \delta^{il}T_{il}$, while the 3 vector degrees of freedom are found by anti-symmetrization $T_i^{(1)} = \frac{1}{2}\varepsilon_{ijl}T_{jl}$. This leaves the traceless symmetric part $T_{il}^{(2)} = \frac{1}{2}(T_{il} + T_{li}) - \frac{1}{3}\delta_{il}T^{(0)}$ to represent the 5 components of the irreducible operator of order 2. The proof of this is analogous to the addition of angular momenta. One first writes the tensor T_{il} in spherical coordinates with $i \to q_1$ and $l \to q_2$. Then the highest component must be $T_2^{(2)} = T_{+1,+1}$ (think of T_{il} as V_iW_l, then $T_2^{(2)} = V_{+1}^{(1)}W_{+1}^{(1)}$). The remaining components $T_q^{(2)}$ are then obtained by acting with J_-.

The Wigner-Eckhart theorem: In a basis $|\alpha, j, m\rangle$ of eigenvectors of J^2 and J_z the matrix elements of an irreducible tensor operator of order k are of the form

$$\boxed{\langle\alpha, j, m|T_q^{(k)}|\alpha', j', m'\rangle = \langle j', m', k, q|j, m,\rangle\langle\alpha, j||T^{(k)}||\alpha' j'\rangle} \tag{9.66}$$

where

- $\langle j', m', k, q | j, m, \rangle$ ... Clebsch-Gordon coefficients (independent of $T_q^{(k)}$),
- $\langle \alpha, j || T^{(k)} || \alpha' j' \rangle$ ... reduced matrix element(independent of $m,\ m',\ q$),
- α ... represents all other quantum numbers.

The Wigner-Eckhart theorem thus factorizes the matrix representation of $T_q^{(k)}$ into a geometric part, which is given by the Clebsch-Gordon coefficients, and a constant, called reduced matrix element, which does not depend on the magnetic quantum numbers.

Proof: For the proof we consider the $(2k+1)(2j'+1)$ vectors

$$T_q^{(k)} |\alpha', j', m'\rangle \qquad (q = -k, ..., k; \quad m' = -j', ..., j') \tag{9.67}$$

and their linear combinations

$$|\sigma, j'', m''\rangle = \sum_{q,m'} T_q^{(k)} |\alpha', j', m'\rangle \langle j', m', k, q | j'', m''\rangle, \tag{9.68}$$

where σ contains j' and α' as well as further quantum numbers that characterize $T^{(k)}$. The crucial point is that the collections $|\sigma, j'', m''\rangle$ of $(2k+1)(2j'+1)$ states for fixed σ and for $|m''| \le j'' \le j' + k$ indeed transform according to the irreducible spin-j'' representations, as is suggested by the notation. This follows from (9.61) and the definitions of tensor operators and Clebsch-Gordon coefficients. Since the latter form a unitary matrix we can invert this transformation and obtain

$$T_q^{(k)} |\alpha', j', m'\rangle = \sum_{j'',m''} |\sigma, j'', m''\rangle \langle j'', m'' | j', m', k, q\rangle. \tag{9.69}$$

If we now multiply this equation with $\langle \alpha, j, m|$ we get

$$\begin{aligned} \langle \alpha, j, m | T_q^{(k)} | \alpha', j', m' \rangle &= \sum_{j'',m''} \langle \alpha, j, m | \sigma, j'', m''\rangle \langle j'', m'' | j', m', k, q\rangle & (9.70) \\ &= \langle \alpha, j, m | \sigma, j, m\rangle \langle j, m | j', m', k, q\rangle & (9.71) \end{aligned}$$

with $\langle \alpha, j, m | \sigma, j, m\rangle \equiv \langle \alpha, j || T^{(k)} || \alpha', j'\rangle$ because $\langle \alpha, j, m | \sigma, j'', m''\rangle$ is 0 except for $j = j''$ and $m = m''$. The scalar product $\langle \alpha, j, m | \sigma, j, m\rangle$ does not depend on m, as can be shown by insertion of $[J_+, J_-] = 2J_3$, and its dependence on j' and $T^{(k)}$ is implicitly contained in σ. □

As an application we consider the spherical harmonics $Y_q^{(k)} \equiv Y_{kq}$ with angular momentum k, operating on wave functions by multiplication,

$$\langle \alpha, j, m | Y_q^{(k)} | \alpha', j', m'\rangle = \delta_{\alpha\alpha'} \langle j', m', k, q | j, m\rangle \langle j || Y_k || j'\rangle. \tag{9.72}$$

The reduced matrix element is

$$\langle j || Y_k || j'\rangle = \langle j, 0, k, 0 | j', 0\rangle \sqrt{\frac{(2j+1)(2k+1)}{(2j'+1)}} \tag{9.73}$$

(see [Messiah] II, appendix C).

9.4 Symmetries of relativistic quantum mechanics

In chapter 7 we introduced the Dirac equation

$$i\hbar\dot{\psi} = H\psi, \qquad H = c\alpha_i(P_i - \frac{e}{c}A_i) + \beta mc^2 + e\phi \tag{9.74}$$

with 4-component spinors ψ and Hermitian 4×4 matrices β and α_i with 2×2 block entries

$$\beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \qquad \alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix} \tag{9.75}$$

which satisfy the anticommutation relations $\{\alpha_i, \alpha_j\} = \delta_{ij}\mathbb{1}_{4\times 4}$, $\{\alpha_i, \beta\} = 0$, and $\beta^2 = \mathbb{1}_{4\times 4}$. In the relativistic notation we distinguish upper and lower indices $\mu = 0, \ldots, 3$ and combine space-time coordinates, scalar and vector potential, and energy-momentum to 4-vectors according to

$$x^\mu = (ct, \vec{x}), \qquad \partial_\mu = (\frac{1}{c}\partial_t, \vec{\nabla}), \qquad A^\mu = (\phi, \vec{A}), \qquad p^\mu = (\frac{1}{c}E, \vec{p}). \tag{9.76}$$

After multiplication with $\beta/\hbar c$ from the left and with the correspondence rule $p_\mu \to i\hbar\partial_\mu$ the Dirac equation (9.74) becomes [5]

$$\left(i\gamma^\mu\partial_\mu - \frac{e}{\hbar c}\gamma^\mu A_\mu - \frac{c}{\hbar}m\right)\psi(t, \vec{x}) = 0, \tag{9.77}$$

where we introduce the four matrices $\gamma^\mu = (\beta, \beta\vec{\alpha})$, which are unitary $(\gamma^\mu)^\dagger = (\gamma^\mu)^{-1}$ and satisfy the Clifford algebra

$$\{\gamma_\mu, \gamma_\nu\} = 2g_{\mu\nu}, \qquad \text{with} \qquad \gamma_\mu = g_{\mu\nu}\gamma^\nu \tag{9.78}$$

so that $(\gamma^0)^2 = \mathbb{1} = -(\gamma^i)^2$ and different γ's anticommute $\gamma^\mu\gamma^\nu = -\gamma^\nu\gamma^\mu$ if $\mu \neq \nu$. The four matrices γ^μ are the relativistic analog of the three Pauli matrices σ_i. Since relativistic spinors have 4 components (describing the spin-up and the spin-down degrees of freedom of particles *and* antiparticles) the γ^μ are 4×4 matrices. Their unitarity implies that only γ^0 is Hermitian while the three matrices γ^i are anti-Hermitian, which can be expressed in the formula

$$(\gamma^\mu)^\dagger = (\gamma^\mu)^{-1} = \gamma^0\gamma^\mu\gamma^0. \tag{9.79}$$

Explicitly the Dirac matrices read

$$\gamma^0 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \qquad \text{[Pauli representation]}. \tag{9.80}$$

Matrix representations γ^μ of the Clifford algebra (9.78) are far from unique, but it can be shown that all unitary representations are related by unitary similarity transformations $\gamma^\mu \to U\gamma^\mu U^{-1}$. For concrete calculations it is usually much better to use the algebraic relations (9.78–9.79) than to use an explicit form of the γ-matrices.[6]

[5] In quantum electrodynamics it is common to introduce Feynman's *slash* notation $\not{a} \equiv \gamma^\mu a_\mu$ [read: a-slash] for contractions of vectors with γ matrices and to set $\hbar = c = 1$ so that the Dirac operator reads $(i\not{\partial} - e\not{A} - m)$ and $\not{a}^2 = a^2\mathbb{1}_{4\times 4}$, which generalizes the nonrelativistic formula $(\vec{v}\vec{\sigma})^2 = v^2\mathbb{1}_{2\times 2}$.

[6] For certain applications particular representations may, however, be useful: The Pauli representation is convenient for taking the non-relativistic limit (see chapter 7). In a Majorana representation all γ^μ are imaginary

$$\gamma^0 = \begin{pmatrix} \sigma_2 & 0 \\ 0 & \sigma_2 \end{pmatrix}, \ \gamma^1 = \begin{pmatrix} i\sigma_1 & 0 \\ 0 & i\sigma_1 \end{pmatrix}, \ \gamma^2 = \begin{pmatrix} 0 & i\sigma_3 \\ i\sigma_3 & 0 \end{pmatrix}, \ \gamma^3 = \begin{pmatrix} i\sigma_3 & 0 \\ 0 & -i\sigma_3 \end{pmatrix} \qquad \text{[Majorana]} \tag{9.81}$$

9.4.1 Lorentz covariance of the Dirac-equation

We want to show now that the Dirac equation (9.77) retains its form under a Lorentz transformation $x'^{\mu} = L^{\mu}{}_{\nu}x^{\nu}$. For given $L^{\mu}{}_{\nu}$ we expect the Dirac spinor ψ to transform linearly $\psi'(x') = \Lambda(L)\psi(x)$ with some 4×4 matrix Λ depending on L. Note that we always use a matrix notation and never write explicit indices for spinors ψ and their linear transformations by multiplications with γ-matrices, which are four 4×4 matrices acting by matrix multiplication on 4-component *spinors* but labeled by a Lorentz vector index μ. One should not be confused by the coincidence that spinors and vectors have the same *number* of components in 4 dimensions. They nevertheless transform differently under Lorentz transformations![7]

For simplicity we consider the free Dirac equation with $A_\mu = 0$. Inserting

$$x'^{\nu} = L^{\nu}{}_{\mu}x^{\mu} \quad \Rightarrow \quad \partial_\mu = \frac{\partial}{\partial x^\mu} = \frac{\partial x'^{\nu}}{\partial x^\mu}\frac{\partial}{\partial x'^{\nu}} = L^{\nu}{}_{\mu}\frac{\partial}{\partial x'^{\nu}} = L^{\nu}{}_{\mu}\partial'_\nu \qquad \text{and} \quad \psi = \Lambda^{-1}\psi' \tag{9.83}$$

into the equation $(i\gamma^\mu\partial_\mu - \frac{c}{\hbar}m)\psi = 0$ we obtain

$$(i\gamma^\mu L^{\nu}{}_{\mu}\partial'_\nu - \tfrac{c}{\hbar}m)\Lambda^{-1}\psi' = 0. \tag{9.84}$$

This transforms into $(i\gamma^\nu\partial'_\nu - \frac{c}{\hbar}m)\psi' = 0$ by multiplication with Λ from the left provided that $\Lambda\gamma^\mu L^{\nu}{}_{\mu}\Lambda^{-1} = \gamma^\nu$ or

$$\boxed{\Lambda^{-1}\gamma^\mu\Lambda = L^{\mu}{}_{\nu}\gamma^\nu.} \tag{9.85}$$

This is the relativistic version of the equation (9.59) and defines the spinor transformation $\Lambda(L)$ in terms of a Lorentz transformation L. Like in the non-relativistic case $\pm\Lambda$ correspond to the same $L^{\mu}{}_{\nu}$ so that the spin group is a double cover of the Lorentz group. The condition (9.85) also guarantees covariance of the interacting Dirac equation (9.77) because the gauge potential A_μ transforms like the gradient ∂_μ.

Finite transformations $\Lambda(L)$ can be obtained by exponentiation of infinitesimal ones,

$$L^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \omega^{\mu}{}_{\nu} + \mathcal{O}(\omega^2) \qquad \Rightarrow \qquad \omega_{\mu\nu} = g_{\mu\rho}\omega^{\rho}{}_{\nu} = -\omega_{\nu\mu}. \tag{9.86}$$

Similarly to the electromagnetic field strength $F_{\mu\nu} = -F_{\nu\mu}$, which contains the electric and the magnetic fields as 3-vectors, the antisymmetric tensor $\omega_{\mu\nu}$ contains infinitesimal spacial

---

so that the free Dirac equation becomes real. Weyl representations are block-offdiagonal

$$\gamma^0 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix} \qquad \text{[Weyl representation]} \tag{9.82}$$

decomposing the massless Dirac equations into two-component equations for left- and right-handed particles.

[7] We already know that spinors in 3 dimensions have 2 compoments, while vectors have 3 components; in higher dimensions $d > 4$, on the other hand, it can be shown that the number of components of spinors grows like $2^{d/2}$, which is much larger than the number d of vector components.

rotations $\delta R_{\vec\rho}$ about an axis $\vec\rho$

$$\delta R_{\vec\rho}\, x^i = \varepsilon_{ijk}\rho^j x^k = \omega^i{}_k x^k \qquad \Rightarrow \qquad \omega_{jk} = \rho^i \varepsilon_{ijk} \tag{9.87}$$

and infinitesimal boosts $\omega_{i0} = v_i/c$ in the direction of a velocity vector $\vec v$, whose finite form for $\vec v = v\vec e_x$ is

$$L^\mu{}_\nu = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad \text{with} \quad \gamma = \frac{1}{\sqrt{1-\beta^2}} \quad \text{and} \quad \beta = \frac{v}{c}. \tag{9.88}$$

With the ansatz $\Lambda(L) = \mathbb{1} + \frac12 \omega_{\mu\nu}\Sigma^{\mu\nu} + \mathcal{O}(\omega^2)$ the defining equation $\Lambda^{-1}\gamma^\rho\Lambda = L^\rho{}_\nu\gamma^\nu$ becomes

$$\tfrac12\,\omega_{\mu\nu}[\gamma^\rho, \Sigma^{\mu\nu}] = \omega^\rho{}_\nu\gamma^\nu = g^{\rho\mu}\omega_{\mu\nu}\gamma^\nu \qquad \Rightarrow \qquad \tfrac12[\gamma_\rho, \Sigma_{\mu\nu}] = \tfrac12(g_{\rho\mu}\gamma_\nu - g_{\rho\nu}\gamma_\mu) \tag{9.89}$$

whose solution can be guessed to be of the form $\Sigma_{\mu\nu} = a[\gamma_\mu, \gamma_\nu]$. Since $[\gamma_\mu,\gamma_\nu] = 2\gamma_\mu\gamma_\nu - 2g_{\mu\nu}\mathbb{1}$ and

$$[\gamma_\rho, \Sigma_{\mu\nu}] = 2a[\gamma_\rho, \gamma_\mu\gamma_\nu] = 2a(\{\gamma_\rho,\gamma_\mu\}\gamma_\nu - \gamma_\mu\{\gamma_\rho,\gamma_\nu\}) = 4a(g_{\rho\mu}\gamma_\nu - \gamma_\mu g_{\rho\nu}) \tag{9.90}$$

equation (9.89) is indeed solved for $a = \frac14$ and the solution can be shown to be unique. Hence

$$\Sigma_{\mu\nu} = \tfrac14[\gamma_\mu, \gamma_\nu]. \tag{9.91}$$

For an alternative derivation of the relativistic spin operator $S_{\mu\nu} = -\frac{\hbar}{i}\Sigma_{\mu\nu}$ we write the spin operator S_i for spacial rotations, which has already been determined in chapter 7, in a Lorentz covariant form. We recall the formula

$$[H, L_i] = -i\hbar c\varepsilon_{ijk}\alpha_j P_k = -[H, S_i], \qquad \text{for} \qquad S_i = \frac{\hbar}{4i}\varepsilon_{ijk}\alpha_j\alpha_k = \frac{\hbar}{2}\begin{pmatrix}\sigma_i & 0 \\ 0 & \sigma_i\end{pmatrix} \tag{9.92}$$

from which we concluded that $\vec J = \vec L + \vec S$ is the conserved angular momentum and $\vec S$ is the spin. With $\vec\rho\vec S = \frac{\hbar}{4i}\rho^i\varepsilon_{ijk}\alpha_j\alpha_k = -\frac{\hbar}{4i}\rho^i\varepsilon_{ijk}\gamma^j\gamma^k$ we conclude that an infinitesimal rotation $\omega_{jk} = \rho^i\varepsilon_{ijk}$ should be given by

$$\delta\psi = -\tfrac{i}{\hbar}\vec\rho\vec S\psi = \tfrac14\omega_{jk}\gamma^j\gamma^k\psi = \tfrac12\omega_{ij}\Sigma^{ij}\psi, \tag{9.93}$$

which indeed is the specialization of our previous result to spacial rotations $\omega_{i0} = 0$, $\omega_{jk} = \rho^i\varepsilon_{ijk}$.

9.4.2 Spin and helicity

A covariant description of the spin of a relativistic particle can be given in terms of the Pauli-Lubanski vector

$$W_\alpha = -\frac12\varepsilon_{\alpha\beta\gamma\delta}J^{\beta\gamma}P^\delta \quad \text{with} \quad \varepsilon^{0123} = 1 = -\varepsilon_{0123} \tag{9.94}$$

where $J^{\mu\nu} = x^\mu P^\nu - x^\nu P^\mu + S^{\mu\nu}$ is the total angular momentum. By evaluation in the center of mass frame it can be shown that the eigenvalues of $W^2 = W^\mu W_\mu$ are

$$W^2 = -m^2c^2\hbar^2\, s(s+1) \tag{9.95}$$

(without proof). For massless particles, however, $W^2 = 0$ so that s cannot be determined by W! The physical reason for this problem is that massless particles can never be in their center of mass frame. In fact, the spin quantum number j refers to the rotation group $SO(3)$ which cannot be used to classify massless particles exactly because of the non-existence of a center of mass frame (indeed, if photons could be described by spin $j = 1$, then the magnetic quantum number would have three allowed values; but we know that photons only have the two tranversal polarizations).

The intrinsic angular momentum of massless particles therefore has to be described by a different conserved quantity. Equation (9.92) implies that $\vec{p} \cdot \vec{\mathcal{S}}$ is a constant of motion

$$[\vec{p} \cdot \vec{\mathcal{S}}, H] = 0. \tag{9.96}$$

If $p = |\vec{p}| \neq 0$, which is always the case for massless particles, we can define the *helicity*

$$s_p = \frac{\vec{p} \cdot \vec{\mathcal{S}}}{p}, \tag{9.97}$$

which is the spin component in the direction of the velocity of the particle. For solutions of the Dirac equation its eigenvalues can be shown to be $s_p = \pm\hbar/2$. For a given momentum a Dirac particle can have two different helicities for positive-energy and two different helicities for negative energy solutions, so that the four degrees of freedom describe particles and antiparticles of both helicities. For the massless neutrinos, however, only positive helicity (left-handed) particles and negative helicity (right-handed) anti-particles exist in the standard model of particle interactions. The massless photons with $s_p = \pm\hbar$ and the gravitons with $s_p = \pm 2\hbar$ are their own antiparticles and they exist with two rather than $2s_p + 1$ polarizations.

9.4.3 Dirac conjugation and Lorentz tensors

If we try to construct a conserved current $j^\mu = (c\rho, \vec{j})$, which satisfies the continuity equation $\dot{\rho} + \mathrm{div}\vec{j} = 0$ and thus generalizes the probability density current of chapter 2, it is natural to consider the quantity $\psi^\dagger \gamma^\mu \psi$, which however is not real

$$(\psi^\dagger \gamma^\mu \psi)^* = (\psi^\dagger \gamma^\mu \psi)^\dagger = \psi^\dagger (\gamma^\mu)^\dagger \psi \neq \psi^\dagger \gamma^\mu \psi \tag{9.98}$$

because the γ^μ is anti-Hermitian for $\mu \neq 0$. It can, hence, also not transform as a 4-vector because Lorentz transformations would mix the real 0-component with the imaginary spacial components.

An appropriate real current can be constructed by replacing the Hermitian conjugate $\psi^\dagger$ by the *Dirac conjugate* spinor

$$\boxed{\overline{\psi} = \psi^\dagger \gamma^0,} \qquad \boxed{j^\mu = \overline{\psi} \gamma^\mu \psi.} \tag{9.99}$$

Now we can use eq. (9.79) and find

$$(j^\mu)^* = (\psi^\dagger\gamma^0\gamma^\mu\psi)^\dagger = \psi^\dagger(\gamma^\mu)^\dagger\gamma^0\psi = \psi^\dagger\gamma^0\gamma^\mu\psi = j^\mu \tag{9.100}$$

so that j^μ is indeed real.

The reason for introducing the Dirac conjugation can also be seen by computing the Lorentz transform of $\psi^\dagger$,

$$\langle\psi'| = \langle\psi|\Lambda^\dagger, \qquad \Lambda^\dagger \neq \Lambda^{-1}. \tag{9.101}$$

Considering infinitesimal transformations we observe that the non-unitarity of Λ again has its origin in the non-Hermiticity γ^μ and thus again can be compensated by conjugation with γ^0.

$$\begin{aligned}\left(\mathbb{1}+\frac{1}{8}\omega_{\mu\nu}[\gamma^\mu,\gamma^\nu]\right)^\dagger &= \mathbb{1}+\frac{1}{8}\omega_{\mu\nu}([\gamma^\mu,\gamma^\nu])^\dagger = \mathbb{1}+\frac{1}{8}\omega_{\mu\nu}[(\gamma^\nu)^\dagger,(\gamma^\mu)^\dagger] = \mathbb{1}-\frac{1}{8}\omega_{\mu\nu}[(\gamma^\mu)^\dagger,(\gamma^\nu)^\dagger]\\ &= \gamma^0(\mathbb{1}-\frac{1}{8}\omega_{\mu\nu}[\gamma^\mu,\gamma^\nu])\gamma^0 = \gamma^0\left(\mathbb{1}+\frac{1}{8}\omega_{\mu\nu}[\gamma^\mu,\gamma^\nu]\right)^{-1}\gamma^0+\mathcal{O}(\omega^2).\end{aligned} \tag{9.102}$$

and hence for finite transformations

$$\boxed{\Lambda^\dagger = \gamma^0\Lambda^{-1}\gamma^0} \tag{9.103}$$

Λ is unitary for purely spacial rotations, but for Lorentz boosts it is not. For the Lorentz transformation of the Dirac adjoint spinor (9.103) implies

$$\psi' = \Lambda\psi \qquad \Rightarrow \qquad \overline{\psi}\,' = \psi^\dagger\Lambda^\dagger\gamma^0 = \psi^\dagger\gamma^0\Lambda^{-1} = \overline{\psi}\Lambda^{-1} \tag{9.104}$$

so that $(\overline{\psi}\psi)' = \overline{\psi}\psi$ is a scalar and the current j^μ transforms as a Lorentz vector,

$$(j^\mu)' = \overline{\psi}\Lambda^{-1}\gamma^\mu\Lambda\psi = L^\mu{}_\nu\overline{\psi}\gamma^\nu\psi = L^\mu{}_\nu j^\nu. \tag{9.105}$$

The divergence of the current $j^\mu = \overline{\psi}\gamma^\mu\psi$ can now be computed using the Dirac equation (9.77) and its conjugate

$$0 = \psi^\dagger\left((-i\overleftarrow{\partial}_\mu - \tfrac{e}{\hbar c}A_\mu)(\gamma^\mu)^\dagger - \tfrac{c}{\hbar}m\right)\gamma^0 = \overline{\psi}\left((-i\overleftarrow{\partial}_\mu - \tfrac{e}{\hbar c}A_\mu)\gamma^\mu) - \tfrac{c}{\hbar}m\right), \tag{9.106}$$

where $\psi^\dagger\overleftarrow{\partial}_\mu \equiv \partial_\mu\psi^\dagger$. For the divergence of the current j^μ we thus obtain

$$\begin{aligned}\partial_\mu(\overline{\psi}\gamma^\mu\psi) &= (\partial_\mu\overline{\psi}\gamma^\mu)\psi + \overline{\psi}(\gamma^\mu\partial_\mu\psi)\\ &= \left(\overline{\psi}(i\tfrac{e}{\hbar c}A_\mu\gamma^\mu + i\tfrac{c}{\hbar}m)\right)\psi + \overline{\psi}\left((-i\tfrac{e}{\hbar c}A_\mu\gamma^\mu - i\tfrac{c}{\hbar}m)\psi\right) = 0,\end{aligned} \tag{9.107}$$

which establishes the continuity equation $\partial_\mu j^\mu = 0$.

Lorentz tensors. Similarly to eq. (9.105) we can compute the Lorentz transformation for the insertion of a product of γ-matrices,

$$(\overline{\psi}\gamma^{\mu_1}\ldots\gamma^{\mu_k}\psi)' = \overline{\psi}\Lambda^{-1}\gamma^{\mu_1}\Lambda\Lambda^{-1}\gamma^{\mu_2}\Lambda\ldots\Lambda^{-1}\gamma^{\mu_k}\Lambda\psi = L^{\mu_1}{}_{\nu_1}\ldots L^{\mu_k}{}_{\nu_k}\overline{\psi}\gamma^{\nu_1}\ldots\gamma^{\nu_k}\psi \tag{9.108}$$

The expectation values $\overline{\psi}\gamma^{\mu_1}\ldots\gamma^{\mu_k}\psi$ hence transform as Lorentz tensors of order k. Since

$$\gamma^\mu\gamma^\nu = \frac{1}{2}[\gamma^\mu,\gamma^\nu] + \frac{1}{2}\{\gamma^\mu,\gamma^\nu\} = \frac{1}{2}[\gamma^\mu,\gamma^\nu] + g^{\mu\nu}\mathbb{1} \tag{9.109}$$

every index symmetrization reduces the number of γ-matrix factors by two so that irreducible tensors are completely antisymmetric in their indices μ_i. In 4 dimensions we can antisymmetrize in at most 4 indices. The complete set of irreducible Lorentz tensors is listed in table 9.1 (we avoid the customary sympols $A^\mu \equiv \widetilde{V}^\mu$ for the axial vector and $P \equiv \widetilde{S}$ for the pseudoscalar to avoid confusion with with gauge potentials and parity).

	Lorentz tensor	$\binom{4}{k}$	C	P	T	CPT
Scalar	$S = \overline{\psi}\psi$	1	S	S	S	S
Vector	$V^\mu = \overline{\psi}\gamma^\mu\psi$	4	$-V^\mu$	V_μ	V_μ	$-V^\mu$
Antisym. tensor	$T^{\mu\nu} = \frac{i}{2}\overline{\psi}[\gamma^\mu,\gamma^\nu]\psi$	6	$-T^{\mu\nu}$	$T_{\mu\nu}$	$-T_{\mu\nu}$	$T^{\mu\nu}$
Axial vector	$\widetilde{V}^\mu = \frac{i}{3!}\varepsilon^{\mu\nu\rho\sigma}\overline{\psi}\gamma_\nu\gamma_\rho\gamma_\sigma\psi$	4	$\widetilde{V}^\mu$	$-\widetilde{V}_\mu$	$\widetilde{V}_\mu$	$-\widetilde{V}^\mu$
Pseudo scalar	$\widetilde{S} = \overline{\psi}\gamma^0\gamma^1\gamma^2\gamma^3\psi$	1	$\widetilde{S}$	$-\widetilde{S}$	$-\widetilde{S}$	$\widetilde{S}$

Table 9.1: Lorentz tensors of order k with $\binom{4}{k}$ components and their CPT transformation.

The total number of components is $\sum_{k=0}^{d}\binom{d}{k} = 2^d = 16$ and it can be shown that the antisymmetrized products of γ matrices are linearly independent.[8] They hence form a basis of all 16 linear operators in the 4-dimensional spinor space. This is the relativistic analog of the fact that the Pauli-matrices together with the unit matrix form a basis for the operators in the 2-dimensional spinor space of nonrelativistic quantum mechanics. The transformation properties under the discrete symmetries C, P and T are indicated in the last columns of table 9.1 and will be discussed in the next section.

9.5 Parity, time reversal and charge-conjugation

The nonrelativistic kinematics is invariant under the discrete symmetries parity $P\vec{x} = -\vec{x}$ and time reversal $Tt = -t$.

Parity. In quantum mechanics the classical action $P\vec{x} = -\vec{x}$ of partity is implemented by a unitary operator $\mathcal{P}$ in Hilbert space transforming $|\psi\rangle \to \mathcal{P}|\psi\rangle$ and hence

$$P\vec{x} = -\vec{x} \qquad \Rightarrow \qquad \mathcal{P}\,\vec{X}\,\mathcal{P}^\dagger = -\vec{X}, \qquad \mathcal{P}\vec{P}\,\mathcal{P}^\dagger = -\vec{P}. \tag{9.110}$$

[8] The proof of linear independence uses the lemma that all products of γ matrices that differ from $\pm\mathbb{1}$ are traceless. A spinor in d dimensions has $2^{[d/2]}$ components ($[d/2]$ is the greatest integer smaller or equal to $d/2$).

While coordinates and momenta are (polar) vectors, i.e. odd under parity, the angular momentum $\vec{x} \times \vec{p}$ is a *pseudo vector*, or *axial vector*), i.e. even under parity

$$L = \vec{x} \times \vec{p} \qquad \Rightarrow \qquad \mathcal{P}\vec{L}\mathcal{P}^\dagger = \vec{L}, \qquad \mathcal{P}\,\vec{S}\,\mathcal{P}^\dagger = \vec{S} \tag{9.111}$$

Electromagnetic and strong interactions, as well as gravity, preserve parity. The form of the Maxwell equations implies that the electric field is a vector while the magnetic field transforms as an axial vector

$$\mathcal{P}\,\vec{E}\,\mathcal{P}^\dagger = -\vec{E}, \qquad \mathcal{P}\,\vec{B}\,\mathcal{P}^\dagger = \vec{B}. \tag{9.112}$$

In the relativistic notation $A^\mu \to A_\mu$, i.e. A^0 is parity even and the vector potential $\vec{A}$ is parity odd. The spin-orbit coupling $\vec{L}\vec{S}$ and the magnetic coupling $\vec{B}(\vec{L}+2\vec{S})$ are axial-axial couplings and hence allowed by parity, while $\vec{E}\vec{S}$ would be a vector-axial coupling and is hence forbidden by parity. The parity of the spherical harmonics is

$$\mathcal{P}|Y_{lm}\rangle = (-1)^l\,|Y_{lm}\rangle \tag{9.113}$$

which is the basis for parity selections rules in atomic physics.

Parity is violated in weak interactions, as was first observed in the radioactive β-decay of polarized ^{60}Co. Since spin is parity-even the emission probability has to be the same for the angles θ and $\pi-\theta$ if parity is conserved. But experiments show that most electrons are emitted opposite to the spin direction $\theta > \pi/2$.

Time reversal. For a real Hamiltonian the effect of an inversion of the time direction can be compensated in the Schrödinger equation by complex conjugation of the wave fuction

$$t \to t' = -t \qquad \Rightarrow \qquad i\hbar\frac{\partial}{\partial t'}\psi^* = i\hbar\frac{\partial}{-\partial t}\psi^* = H\psi^*. \tag{9.114}$$

Time reversal therefore is implemented in quantum mechanics by an *anti*-unitary operator

$$\mathcal{T}\psi(t,\vec{x}) = \psi^*(-t,\vec{x}) \qquad \Rightarrow \qquad \langle\mathcal{T}\varphi|\mathcal{T}\psi\rangle = \langle\varphi|\psi\rangle^*, \qquad \mathcal{T}\alpha|\psi\rangle = \alpha^*\mathcal{T}|\psi\rangle \tag{9.115}$$

which implies complex conjugation of scalar products but leaves the norms $\sqrt{\langle\psi|\psi\rangle}$ invariant. For antilinear operators Hermitian conjugation can be defined by $\langle\varphi|\mathcal{T}^\dagger|\psi\rangle = \langle\psi|\mathcal{T}|\varphi\rangle$. Anti-unitary is then equivalent to antilinearity and $\mathcal{T}\mathcal{T}^\dagger = 1$. Since velocities and momenta change their signs under time inversion we have

$$\mathcal{T}\,X_i\,\mathcal{T}^{-1} = X_i, \qquad \mathcal{T}\,P_i\,\mathcal{T}^{-1} = -P_i \tag{9.116}$$

The above formulas are compatible with the canonical commutation relations

$$\mathcal{T}[P_i,X_j]\mathcal{T}^{-1} = [-P_i,X_j] = -\frac{\hbar}{i}\delta_{ij} = \mathcal{T}\frac{\hbar}{i}\delta_{ij}\mathcal{T}^{-1}. \tag{9.117}$$

Invariance of the Maxwell equations under time reversal implies

$$\mathcal{T}\,\vec{E}\,\mathcal{T}^{-1}=\vec{E},\qquad \mathcal{T}\,\vec{B}\,\mathcal{T}^{-1}=-\vec{B},\qquad \mathcal{T}\,A^0\,\mathcal{T}^{-1}=A^0,\qquad \mathcal{T}\,\vec{A}\,\mathcal{T}^{-1}=-\vec{A} \tag{9.118}$$

so that gauge potentials transforms in the same way $A^\mu \to A_\mu$ as under parity. In fundamental interactions violation of time reversal invariance has only been observed for weak interactions.

9.5.1 Discrete symmetries of the Dirac equation

In addition to parity and time reversal the relativistic theory has another discrete symmetry, called charge conjugation, which is the exchange of particles and anti-particles.

Parity. Invariance of the Dirac equation under the parity transformation $\vec{x} \to -\vec{x}$ implies that $(i\gamma^\mu \partial'_\mu - m)\psi'(x') = 0$ should be equivalent to $(i\gamma^\mu\partial_\mu - m)\psi(x) = 0$ for $\psi' = \mathcal{P}\psi$, hence

$$\mathcal{P}^{-1}\left(i\left(\gamma^0\frac{\partial}{\partial x^0}+\gamma^i\frac{\partial}{\partial(-x^i)}\right)-m\right)\mathcal{P}\psi=(i\gamma^\mu\frac{\partial}{\partial x^\mu}-m)\psi \tag{9.119}$$

which implies

$$\mathcal{P}^{-1}\gamma^0\mathcal{P}=\gamma^0,\quad \mathcal{P}^{-1}\gamma^i\mathcal{P}=-\gamma^i \qquad\Rightarrow\qquad \mathcal{P}\,|\psi\rangle=\gamma^0\,|\psi\rangle. \tag{9.120}$$

Equation (9.119) actually fixes the action of the unitary parity operator $\mathcal{P}$ on spinors only up to an irrelevant phase factor and we followed the usual choice.

Charge conjugation. According to Dirac's hole theory exchange of particles and anti-particles should reverse the signs of electric charges and hence the sign of the gauge potential

$$\mathcal{C}\,\vec{E}\,\mathcal{C}^{-1}=-\vec{E},\qquad \mathcal{C}\,\vec{B}\,\mathcal{C}^{-1}=-\vec{B},\qquad \mathcal{C}\,A^\mu\,\mathcal{C}^{-1}=-A^\mu. \tag{9.121}$$

The derivation of the action of $\mathcal{C}$ on spinor starts with the observation that the relative sign between the kinetic and the gauge term is reversed in the conjugated Dirac equation (9.106). Transposition of that equation yields

$$\left((-\gamma^\mu)^T(i\partial_\mu+\tfrac{e}{\hbar c}A_\mu)-\tfrac{c}{\hbar}m\right)\overline{\psi}^T=0. \tag{9.122}$$

This lead to the condition

$$\mathcal{C}^{-1}\,(-\gamma^\mu)^T\,\mathcal{C}=\gamma^\mu \qquad\Rightarrow\qquad \mathcal{C}\,\psi=i\gamma^2\gamma^0\,\overline{\psi}^T=i\gamma^2\psi^* \tag{9.123}$$

in the standard representation[9] (9.80). Charge conjugation is hence an anti-unitary operation. Since $i\gamma^2=\begin{pmatrix}0 & i\sigma_2\\ -i\sigma_2 & 0\end{pmatrix}=\begin{pmatrix}0 & -\varepsilon\\ \varepsilon & 0\end{pmatrix}$ with $\varepsilon=\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$ charge conjugation exchanges positive and negative energy solutions and their chiralities.

[9] While $\mathcal{P}=\gamma^0$ is representation independent the explicit form of the antilinear operators $\mathcal{C}$ and $\mathcal{T}$ in terms of γ-matrices depends on the representation used for

Time reversal. The complex conjugate Dirac equation for $t' = -t$ is $(-i\gamma^{\mu *}\partial'_\mu - m)\psi^* = 0$. With the ansatz $\mathcal{T}|\psi\rangle = B|\psi^*\rangle$ this implies

$$B^{-1}(-i\gamma^{\mu *})B = i\gamma^0\gamma^\mu\gamma^0 \qquad \Rightarrow \qquad B = i\gamma^1\gamma^3, \qquad \mathcal{T}|\psi\rangle = i\gamma^1\gamma^3|\psi^*\rangle \tag{9.124}$$

in the standard representations (with the customary choice of the phase).

The transformation properties for the Lorentz tensors that follow from the formulas (9.120, 9.123, 9.124) for the $\mathcal{P}$, $\mathcal{C}$ and $\mathcal{T}$ are listed in table 9.1. The CPT theorem states that the combination of these three discrete transformations is a symmetry in every local Lorentz-invariant quantum field theory. The proof is based in the fact that all Lorentz tensors of order k (the complete set of fermion bilinears in table 9.1 as well as scalar fields and gauge fields A^μ) transform with a factor $(-1)^k$ under CPT and that Lorentz invariant interaction terms have no free Lorentz indices. Violation of time reversal invariance thus becomes equivalent to CP violation, which was first observed in 1964.[10]

9.6 Gauge invariance and the Aharonov–Bohm effect

An important aspect of the electromagnetic interaction with quantum particles is the fact that the interaction term in the Dirac equation

$$(\gamma^\mu(i\partial_\mu - \tfrac{e}{\hbar c}A_\mu) - \frac{c}{\hbar}m)\psi = 0, \tag{9.127}$$

[10] CP-violation in kaon decay is observed as follows [Nachtmann]. The theory of strong interactions implies that nucleons like the proton $|p\rangle = |uud\rangle$ with $m_p = 938MeV$ and the neutron $|n\rangle = |udd\rangle$ with $m_n = 940MeV$ consist of three quarks, while mesons like the pions $|\pi^+\rangle = |u\bar{d}\rangle$, $|\pi^-\rangle = |d\bar{u}\rangle$ with $m_{\pi^\pm} = 140MeV$ and $|\pi^0\rangle = (|u\bar{u}\rangle - |d\bar{d}\rangle)/\sqrt{2}$ with $m_{\pi^0} = 135MeV$ consist of a quark and an anti-quark. The K mesons $|K^+\rangle = |u\bar{s}\rangle$, $|K^-\rangle = |s\bar{u}\rangle$ = with $m_{K^\pm} = 494MeV$ and $|K^0\rangle = |d\bar{s}\rangle$, $|\overline{K}^0\rangle = |s\bar{d}\rangle$ with $m_{K^0} = 498MeV$ contain the somewhat heavier strange quark s and hence can decay by weak interactions.

Now the states $|K^0\rangle = \mathcal{C}|\overline{K}^0\rangle$ and $|\overline{K}^0\rangle = \mathcal{C}|K^0\rangle$ are each others antiparticles and both are observed to be parity odd $\mathcal{P}|K^0\rangle = -|K^0\rangle$ (according to the parity conserving strong processes in which they are created). The neutral K-mesons can only decay by weak interactions,

$$|K^0\rangle,\ |\overline{K}^0\rangle \quad \to \quad \pi^+\pi^-,\ \pi^0\pi^0, \quad \pi^+\pi^-\pi^0,\ \pi^0\pi^0\pi^0,\ \pi^\pm e^\mp\nu,\ \pi^\pm\mu^\mp\nu. \tag{9.125}$$

which break parity as well as charge conjugation. If we assume that weak interactions preserve the combination CP then the CP eigenstates

$$|K^0_{(\pm)}\rangle = \tfrac{1}{\sqrt{2}}(|K^0\rangle \mp |\overline{K}^0\rangle, \quad \mathcal{CP}|K^0_{(\pm)}\rangle = \pm|K^0_{(\pm)}\rangle \quad \to \quad |K_S\rangle \equiv |K^0_{(+)}\rangle, \quad |K_L\rangle \equiv |K^0_{(-)}\rangle \tag{9.126}$$

can only decay into CP eigenstates. In particular, $|K_S\rangle$ (S=short lived, with $\tau_s = 0.89 \cdot 10^{-10}s$) can decay into two pions (a CP-even state because $l = 0$ by angular momentum conservation), while $|K_L\rangle$ (L=long lived, with $\tau_L = 5.18 \cdot 10^{-8}s$) can only have the less likely 3-particle decays. But $|K_L\rangle$ is observed to decay into two pions with a probability of about 0.3%. Moreover, the 3-particle decay of $|K_L\rangle$ producing a positively charged lepton is observed to be 0.66% more likely than its decay into the CP conjugate states containing an electron e^- or a muon μ^-.

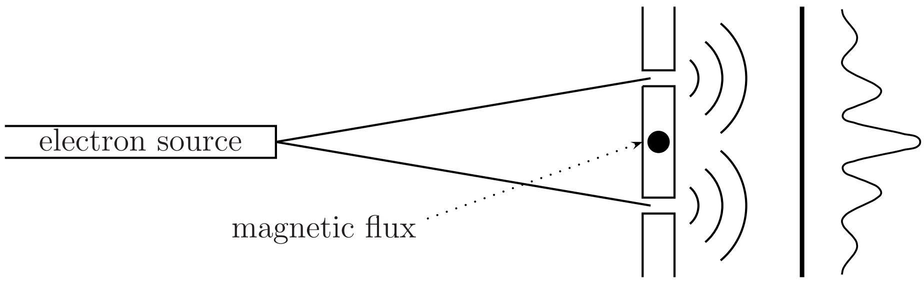

Figure 9.1: In the modified double slit experiment proposed by Aharonov and Bohm one observes a shift of the interference pattern proportional to the magnetic flux although the electrons only move in field-free regions.

like in the Schrödinger equation, explicitly depends on the gauge potential A_μ, which is not observable because the electromagnetic fields $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ are invariant under $A_\mu \to A'_\mu = A_\mu - \partial_\mu \Lambda$. We already noticed in chapter 2 that the complete wave equation is invariant under this gauge transformation if we simultaneoly change the equally unobservable phase of the wave function

$$A'_\mu = A_\mu - \partial_\mu \Lambda(t, \vec{x}) \qquad \Rightarrow \qquad \psi' = e^{\frac{ie}{\hbar c}\Lambda}\psi \tag{9.128}$$

because

$$(i\partial_\mu - \tfrac{e}{\hbar c}A'_\mu)e^{\frac{ie}{\hbar c}\Lambda} = e^{\frac{ie}{\hbar c}\Lambda}(i\partial_\mu - \tfrac{e}{\hbar c}A_\mu). \tag{9.129}$$

In the nonrelativistic limit this gauge invariance, split into space and time components, is inherited to the Schrödinger equation.

In 1959 Y. Aharonov and D. Bohm made the amazing prediction of an apparent action at a distance due to the form of the gauge interaction, which was experimentally verified by R.C. Chambers in 1960. This phenomenon is also used for practical applications like SQIDS (superconducting quantum interference devices) and it implies flux quantization in superconductors [Schwabl].

The experimental setup is a modification of the double slit experiment as shown in figure 9.1 where a magnetic flux is put between the two electron beams. For an infinitely long coil the B-field is confined inside the coil so that the flux lines cannot reach the domain where the electrons move. Nevertheless, a flux $\phi_B = \int \vec{B} d\vec{S}$ between the two rays leads to a relative phase shift

$$\delta(\arg(\psi_1) - \arg(\psi_2)) = \frac{e}{\hbar c}\phi_B \tag{9.130}$$

and hence to a shift in the interference pattern on the screen behind the slits.

In the Aharonov–Bohm experiment all electromagnetic fields are static. In a domain without magnetic flux $\vec{B} = \vec{\nabla} \times \vec{A} = 0$ the vector potential $\vec{A}$ is curl-free and hence can locally (in a contractible domain) be written as a divergence $\vec{A} = \vec{\nabla}\Lambda(\vec{x})$. This can be considered as a gauge transform of the special solution $\vec{A} = 0$. The "potential" Λ for the vector potential $\vec{A} = \text{grad}\Lambda$ can be computed as a line integral $\Lambda(x) = \int_{x_0}^{x} \vec{A} \cdot d\vec{s}$, which is invariant under continuous deformations of the path as long as we stay in regions without magnetic flux. The computation of the phase shift (9.130) now uses this fact to relate the wave functions $\psi_i^{(0)}$ of the coherent electron beams in the double slit experiment without magnetic field, which are solutions to the Schrödinger equation with $\vec{A} = 0$, to the wave functions $\psi_i^{(B)}$ by gauge transformations. For each of the two beams the trajectories C_i belong to a contractible domain so that we can choose a gauge

$$\Lambda_i(\vec{x}) = \int_{C_i(x)} \vec{A} \cdot d\vec{s} \qquad \Rightarrow \qquad \psi_i^{(B)}(\vec{x}) = \psi_i^{(0)}(\vec{x})\, e^{\frac{ie}{\hbar c}\Lambda_i(\vec{x})}, \tag{9.131}$$

where the contour $C_i(x)$ starts at the electron source and extends along C_i to the position of the electron. Since the initial points of the paths C_i at the electron source and their final points at the screen where the interference is observed are the same for both paths the difference between the phase shifts is

$$\frac{e}{\hbar c}\left(\int_{C_2} \vec{A} \cdot d\vec{s} - \int_{C_1} \vec{A} \cdot d\vec{s}\right) = \frac{e}{\hbar c}\oint \vec{A}\cdot d\vec{s} = \frac{e}{\hbar c}\int (\nabla \times \vec{A})\, d\vec{S} = \frac{e}{\hbar c}\int \vec{B}\, d\vec{S} = \frac{e}{\hbar c}\phi_B, \tag{9.132}$$

where Stockes' theorem has been used to convert the circle integral extending along the closed path $C_2 - C_1$ into a surface integral over a surface enclosed by the beams. This surface integral $\int (\nabla \times \vec{A}) d\vec{S}$ measures the complete magnetic flux between the beams. This completes the derivation of the phase shift (9.130). Since the gauge transformation is the same for the Dirac equation and for the Schrödinger equation the relativistic and the nonrelativistic calculations are identical.