Symmetries

Generalia

Most generally, a symmetry of a physical system is a transform of system's states or configurations

$$A : \mathcal{S} \mapsto \mathcal{S}' \tag{1}$$

that turns a state $\mathcal{S}$ into a physically similar state $\mathcal{S}'$ with a similar dynamics. The symmetries of any system always form a group where the group product is consecutive action:

$$\begin{tikzcd} \boxed{\mathcal{S}''} & \boxed{\mathcal{S}'} \arrow[l, "B"'] & \boxed{\mathcal{S}} \arrow[l, "A"'] \arrow[ll, bend left=50, "BA"] \end{tikzcd} \tag{2}$$

Indeed, this product is always associative,

$$\begin{tikzcd} \boxed{\mathcal{S}'''} & \boxed{\mathcal{S}''} \arrow[l, "C"'] & \boxed{\mathcal{S}'} \arrow[l, "B"'] & \boxed{\mathcal{S}} \arrow[l, "A"'] \arrow[lll, bend left=40, "(CB)A = C(BA)"] \end{tikzcd} \tag{3}$$

the trivial transform $1 : \mathcal{S} \mapsto$ same $\mathcal{S}$ acts as the unity element, $1 \times A = A \times 1 = A$, and every symmetry transform is invertible

$$\begin{tikzcd} \boxed{\mathcal{S}} \arrow[r, shift left, "A^{-1}"] & \boxed{\mathcal{S}'} \arrow[l, shift left, "A"] \end{tikzcd} \qquad AA^{-1} \;=\; A^{-1}A \;=\; 1. \tag{4}$$

Together, these features define a group. However, depending on the specifics of the symmetries in question, the symmetry group cab be either be abelian (commutative product, $AB = BA$ for any A and B) or non-abelian (non-commutative product, $AB \neq BA$ for some A and B).

Symmetries of Classical Field Theories

The state of a classical field theory is the field configuration $\Phi_a(x)$ (for all x and all a), so the symmetries act on such states as

$$A : \Phi_a(x) \mapsto \Phi'_a(x) \quad \text{which is a function of the } \Phi_a(x), \text{ other } \Phi_b(x), \text{ and maybe also } \Phi_b(y) \text{ at other } y \neq x. \tag{5}$$

Example: spacetime translations,

$$\Phi'_a(x) \;=\; \Phi_a(x+d) \qquad \text{for a fixed vector } d^\mu. \tag{6}$$

Example: phase symmetry of a complex scalar field $\Phi(x)$,

$$\Phi'(x) \;=\; e^{-i\theta} \times \Phi(x), \quad \Phi^{*\prime}(x) \;=\; e^{+i\theta} \times \Phi^*(x). \tag{7}$$

A symmetry of a field theory must preserve the fields' equations of motion. That is, *if* the original fields $\Phi_a(x)$ obey their EoM (equations of motion), *then* the symmetry-transformed fields $\Phi'_a(x)$ must also obey their EoM, *and vice verse.* For example, for the self-interacting complex field Φ and its conjugate, their EoM

$$(\partial^2 + m^2)\Phi \;=\; -\lambda|\Phi|^2 \times \Phi, \qquad (\partial^2 + m^2)\Phi^* \;=\; -\lambda|\Phi|^2 \times \Phi^*, \tag{8}$$

are invariant under the phase symmetry (7).

Usually *the symmetries preserve the action of the theory*:

$$S[\Phi_a(x)] \;=\; \int d^4x\, \mathcal{L}(\text{all}\,\Phi_a,\, \text{all}\,\partial_\mu\Phi_a), \tag{9}$$

thus

$$S[\Phi'_a(x)] \;=\; S[\Phi_s(a)]. \tag{10}$$

For the so-called *on-shell symmetries* the action is preserved only when all the fields obey their EoM. But the *off-shell symmetries* preserve the action of any field configuration, even if it does not obey any EoM. For example, consider the phase symmetry (7) of the complex scalar field with the Lagrangian density

$$\mathcal{L}\bigl(\Phi,\Phi^*,\partial\Phi,\partial\Phi^*\bigr)\ =\ \partial_\mu\Phi^*\,\partial^\mu\Phi\ -\ m^2\Phi^*\Phi\ -\ \frac{\lambda}{2}(\Phi^*\Phi)^2, \tag{11}$$

hence the Euler–Lagrange EoM (8). Clearly, every term in the Lagrangian (11) is invariant under the phase symmetry

$$\Phi'(x)\ =\ e^{-i\theta}\times\Phi(x),\quad \Phi^{*\prime}(x)\ =\ e^{+i\theta}\times\Phi^*(x),\quad \text{same }\theta\ \forall x, \tag{7}$$

regardless of the EoM. That is, for any configuration of the $\Phi(x)$ and $\Phi^*(x)$ fields — whether they obey their EoM or don't obey, — we have

$$\mathcal{L}\bigl(\Phi',\Phi^{*\prime},\partial\Phi',\partial\Phi^{*\prime}\bigr)\ =\ \mathcal{L}\bigl(\Phi,\Phi^*,\partial\Phi,\partial\Phi^*\bigr) \tag{12}$$

and hence

$$S[\Phi',\Phi^{*\prime}]\ =\ S[\Phi,\Phi*]. \tag{13}$$

Thus, the phase symmetry (7) is an off-shell symmetry.

From now on — till the end of the whole course — we shall focus on the off-shell symmetries.

Continuous versus Discrete Symmetries.

Examples of continuous symmetries of field theories:

- The phase symmetry (7).
- Translations of space and time.
- Rotations of space.
- Continuous Lorentz symmetries.
- The isospin symmetry.

— *etc.*, *etc.*

Examples of discrete symmetries:

- Parity AKA space reflection $P : \mathbf{x} \mapsto -\mathbf{x}$ (but $P : t \mapsto +t$).
- Time reversal $T : t \mapsto -t$ (but $T : \mathbf{x} \mapsto +\mathbf{x}$).
- Charge conjugation C : particles $\leftrightarrow$ antiparticles.
  (Also, for the EM field $C : A^\mu(x) \mapsto -A^\mu(x)$).
- $D : \Phi(x) \mapsto -\Phi(x)$ for the real scalar field $\Phi(x)$ with $\lambda\Phi^4$ self-interaction.

— *etc., etc.*

Spacetime versus Internal Symmetries.

The spacetime symmetries relate the new fields $\Phi_a(x)$ to the old fields at different locations $y \neq x$. The continuous spacetime symmetries comprise the translations and the (continuous) Lorentz symmetries

$$\begin{aligned} \text{translation} : \Phi'_a(x) &= \Phi_a(x+d) \quad \text{for a fixed vector } d^\mu, && (14) \\ \text{Lorentz} : \Phi'_a(x) &= \Phi_a(Lx) \quad \text{where } (Lx)^\mu = L^\mu{}_\nu x^\nu, \quad L \in SO^+(3,1), && (15) \end{aligned}$$

Together, the continuous Lorentz symmetries and the translations comprise the *Poincaré group*. There are also discrete spacetime symmetries P (parity) and T (time reversal).

The internal symmetries relate the new $\Phi'_a(x)$ to the old $\Phi_b(x)$ at the same location x. For example, the phase symmetry

$$\Phi'(x) = e^{-i\theta} \times \Phi(x), \quad \Phi^{*\prime}(x) = e^{+i\theta} \times \Phi^*(x), \tag{7}$$

of a complex scalar field is an internal symmetry. For another example, consider the $O(N)$ internal symmetry of N real scalar fields $\Phi_a(x)$ ($a = 1, \ldots, N$) with the Lagrangian

$$\mathcal{L} = \frac{1}{2} \sum_{a=1}^{N} \partial_\mu \Phi_a \partial^\mu \Phi_a - \frac{m^2}{2} \sum_a \Phi_a^2 - \frac{\lambda}{4} \left(\sum_a \Phi_a^2 \right)^2 . \tag{16}$$

This Lagrangian — and hence the action — has an $O(N)$ group of symmetries which act as 'rotations' (and reflections) in the N-dimensional field space. That is,

$$\Phi'_a(x) \;=\; \sum_b R_{ab} \times \Phi_b(x) \tag{17}$$

where $\|R_{ab}\|$ is a real orthogonal $N \times N$ matrix. The orthogonality means $R^\top R = RR^\top = 1$ and hence

$$\sum_a \bigl(\Phi'_a(x)\bigr)^2 \;=\; \sum_b \bigl(\Phi_a(x)\bigr)^2, \tag{18}$$

which preserves the potential terms in the Lagrangian (16). As to the kinetic term in the Lagrangian, we need to use the same matrix R at every spacetime point x. This leads to

$$\partial_\mu\bigl(R_{ab}\Phi_b(x)\bigr) \;=\; \cancel{(\partial_\mu R_{ab})}\Phi_b(x) \;+\; R_{ab}\partial_\mu\Phi_b(x) \;=\; R_{ab}\partial_\mu\Phi_b(x), \tag{19}$$

hence

$$\partial_\mu\Phi'_a(x) \;=\; \sum_b R_{ab}\partial_\mu\Phi_b(x), \tag{20}$$

and therefore

$$\sum_a \bigl(\partial_\mu\Phi'_a(x)\bigr)^2 \;=\; \sum_a \bigl(\partial_\mu\Phi_a(x)\bigr)^2. \tag{21}$$

Thus, the kinetic term in the Lagrangian (16) is indeed preserved by the $O(N)$ symmetries.

In the quantum field theory, the Coleman–Mandula theorem says that a continuous symmetry must be either spacetime or internal but not a mixture of the two types. In the group theory terms, this means that net group of continuous symmetries of some field theory is the *direct product* of the Poincare group of the spacetime symmetries and the group of purely internal symmetries of the theory,

$$G(\text{continuous}) \;=\; G(\text{Poincare}) \otimes G(\text{internal}). \tag{22}$$

However, this theorem does not apply to the discrete symmetries. Indeed, there are important QFTs which do not have separate P (space reflection) or C (charge conjugation) symmetries but do have a combined CP symmetry (simultaneous space reflection and charge conjugation).

Also, the Coleman–Mandula theorem is limited to the bosonic symmetries parametrized by the ordinary real or complex numbers, but does not work for the fermionic symmetries parametrized by the anticommuting Grassmann numbers. In particular, the *supersymmetry* has both internal and spacetime aspects.

Global versus Local Symmetries.

Under a global phase symmetry, a charged field $\Phi(x)$ transforms to

$$\Phi'(x) \;=\; \exp(-i\theta)\times\Phi(x) \quad \text{for the same } \theta \text{ at all } x. \tag{23}$$

But under a local phase symmetry, the $\Phi(x)$ transforms to

$$\Phi'(x) \;=\; \exp(-i\theta(x))\times\Phi(x) \quad \text{for independent } \theta(x) \text{ at each } x. \tag{24}$$

Likewise, in the global $SO(N)$ symmetry of N real fields we have the same $\|R_{ab}\|$ matrix at all x, but in a local $SO(N)$ symmetry we would have independent $\|R_{ab}\|(x)$ matrices at each x.

I shall explain the local symmetries in great detail in later classes. For the moment, let me simply mention the terminological problem when physicists talk to mathematicians. Alas, what the physicists call a global symmetry (because it's the same symmetry all over the spacetime) the mathematicians call a local symmetry (because once you know how it acts at one point, you know everything about it). And vice verse, what the physicists call a local symmetry (because it's action is different at each location x), the mathematicians call a global symmetry (because it has global topological aspects). Fortunately, there are common terms to both physicists and mathematicians: the *rigid symmetries* and the *gauge symmetries* (or gauged symmetries). Here is the dictionary table:

Physics	Math	Common
global	local	rigid
local	global	gauge

Symmetries in Quantum Mechanics

In quantum mechanics, the symmetries act as *unitary operators in the Hilbert space* of the quantum system.⋆ In the Schrödinger picture of quantum mechanics, the symmetry operators act on the quantum states while the operators remain unchanged,

$$\left|\psi'\right\rangle \;=\; \hat{U}\left|\psi\right\rangle, \qquad \hat{A}' \;=\; \hat{A}, \tag{25}$$

while in the Heisenberg picture the symmetries act on the operators while the states remain unchanged,

$$\left|\psi'\right\rangle \;=\; \left|\psi\right\rangle, \qquad \hat{A}' \;=\; \hat{U}^\dagger \hat{A}\hat{U}. \tag{26}$$

Either way, the matrix elements of operators transform as

$$\langle\chi|\,\hat{A}\,|\Psi\rangle \;\mapsto\; \left\langle\chi'\right|\hat{A}'\left|\psi'\right\rangle \;=\; \langle\chi|\,\hat{U}^\dagger\hat{A}\hat{U}\,|\psi\rangle\,. \tag{27}$$

The unitary operators $\hat{U}$ form a *representation* of the symmetry group G. That is, for every symmetry in the group there is a corresponding unitary operator $\hat{U}(g)$, and the map $g \mapsto \hat{U}(g)$ respects the group product: for any two symmetries $g_1, g_2 \in G$

$$\hat{U}(\text{group product } g_2 g_1) \;=\; \text{operator product } \hat{U}(g_2)\hat{U}(g_1). \tag{28}$$

Likewise, $\hat{U}(1) = \hat{1}$ (the unit operator) and

$$\hat{U}(g^{-1}) \;=\; \hat{U}^{-1}(g) \;=\; \hat{U}^\dagger(g). \tag{29}$$

For example, consider the rotation symmetry in the 3D space. Classically, the rotations

⋆ The time reversal symmetry is an exception: Unlike the other symmetries, it acts as an anti-unitary rather than unitary operator. Please see §4.4 of the Sakurai's *Modern Quantum Mechanics* for details.

act as $SO(3)$ matrices,

$$x'_i \;=\; R_{ij}x_j \quad \text{for } R\in SO(3). \tag{30}$$

For a rotation through angle ϕ around an axis in the direction of a unit vector $\mathbf{n}$,

$$R(\mathbf{n},\phi) \;=\; \cos\phi\times\delta_{ij} \;+\; \sin\phi\times\epsilon_{ikj}n_k \;+\; (1-\cos\phi)\times n_in_j\,. \tag{31}$$

In the Hilbert space of a quantum theory, the rotation matrices are represented by the operators

$$\hat{\mathcal{D}}(R(\mathbf{n},\phi)) \;=\; \hat{\mathcal{D}}(\mathbf{n},\phi) \;=\; \exp\bigl(-i\phi\mathbf{n}\cdot\hat{\mathbf{J}}\bigr) \tag{32}$$

(in $\hbar=1$ units) where $\hat{\mathbf{J}}$ is the net angular momentum operator. In group theory terms, the 3 components $\hat{J}_x$, $\hat{J}_y$, $\hat{J}_z$ of the angular momentum *generate* the rotation group, or rather generate its representation in the Hilbert space of the quantum system.

To show that the unitary operators $\hat{\mathcal{D}}(R)$ indeed represent the rotations R, let me start with the rotations through infinitesimal angles ϕ,

$$R_{ij}(\mathbf{n},\,\text{infinitesimal}\,\phi) \;=\; \delta_{ij} \;+\; \phi n_k\epsilon_{ikj} \;+\; O(\phi^2). \tag{33}$$

For convenience, let's define antisymmetric imaginary (and hence hermitian) 3×3 matrices $t_{x,y,z}$ as

$$(t_k)_{ij} \;\stackrel{\text{def}}{=}\; i\epsilon_{ikj} \tag{34}$$

which obey the angular-momentum-like commutation relations

$$[t_k,t_\ell] \;=\; i\epsilon_{k\ell m}t_m\,. \tag{35}$$

In terms of these matrices, eq. (33) becomes

$$R(\mathbf{n},\,\text{infinitesimal}\,\phi) \;=\; 1 \;-\; i\phi n_kt_k \;+\; O(\phi^2). \tag{36}$$

Next, a rotation through a large angle obtains as a large number of consecutive rotations through small angles,

$$R(\mathbf{n},\phi) \;=\; \left[R(\mathbf{n},\tfrac{\phi}{N})\right]^N. \tag{37}$$

For any finite angle ϕ, the $N\to\infty$ limit of $\frac{\phi}{N}$ is infinitesimal, so a rotation through that finite angle obtains as

$$\begin{aligned} R(\mathbf{n},\phi) \;&=\; \lim_{N\to\infty}\left[R(\mathbf{n},\tfrac{\phi}{N})\right]^N \\ &=\; \lim_{N\to\infty}\left[1\,-\,\frac{i\phi n_k t_k}{N}\,+\,O\left(\frac{\phi^2}{N^2}\right)\right]^N \\ &=\; \text{matrix }\exp\bigl(-i\phi n_k t_k\bigr). \end{aligned} \tag{38}$$

Note similarity of this formula to eq. (32) for the $\hat{\mathcal{D}}(\mathbf{n},\phi)$ operators.

Next, consider a product of two finite rotations,

$$R(\mathbf{n}_2,\phi_2)R(\mathbf{n}_1,\phi_1) \;=\; R(\mathbf{n}_3,\phi_3) \quad \text{for some } \mathbf{n_3},\phi_3\,. \tag{39}$$

Since each rotation matrix here is a matrix exponential (38), we may use the Campbell–Baker–Hausdorff formula: for the products of matrix or operator exponentials:

$$\exp(A)\exp(B) \;=\; \exp(C) \quad \text{ for} \tag{40}$$

$$\begin{aligned} C \;&=\; A\,+\,B\,+\,\tfrac{1}{2}[A,B]\,+\,\tfrac{1}{12}[(A-B),[A,B]] \\ &\qquad +\;\text{power series in higher multiple commutators.} \end{aligned} \tag{41}$$

In particular, for the rotation matrices (38) this formula yields

$$\begin{aligned} -i\phi_3 n_{3k}t_k \;&=\; -i(\phi_2 n_{2k}\,+\,\phi_1 n_{1k})t_k\,-\,\frac{1}{2}[\phi_2 n_{2\ell}t_\ell,\phi_1 n_{1m}t_m] \\ &\qquad +\;\frac{i}{12}\bigl[(\phi_2 n_{2\ell}-\phi_1 n_{1\ell})t_\ell,[\phi_2 n_{2m}t_m,\phi_1 n_{1p}t_p]\bigr]\;+\;\cdots \\ &\qquad \langle\!\langle\,\text{using the commutation relations (35)}\,\rangle\!\rangle \\ &=\; -i(\phi_2 n_{2k}\,+\,\phi_1 n_{1k})t_k\,-\,\frac{1}{2}\phi_2 n_{2\ell}\phi_1 n_{1m}\times i\epsilon_{\ell mk}t_k \\ &\qquad +\;\frac{i}{12}(\phi_2 n_{2\ell}-\phi_1 n_{1\ell})\phi_2 n_{2m}\phi_1 n_{1p}\times i\epsilon_{mpq}i\epsilon_{\ell qk}t_k\;+\;\cdots, \end{aligned} \tag{42}$$

which indeed obtains for

$$\begin{aligned}\phi_3 n_{3k} &= (\phi_2 n_{2k} + \phi_1 n_{1k}) + \tfrac{1}{2}\epsilon_{k\ell m}\phi_2 n_{2\ell}\phi_1 n_{1m} \\ &\quad + \tfrac{1}{12}\epsilon_{k\ell q}(\phi_2 n_{2\ell} - \phi_1 n_{1\ell}) \times \epsilon_{qmp}\phi_2 n_{2m}\phi_1 n_{1p} + \cdots\end{aligned} \tag{43}$$

or in vector notations

$$\begin{aligned}(\phi_3 \mathbf{n}_3) &= (\phi_2 \mathbf{n}_2) + (\phi_1 \mathbf{n}_1) + \tfrac{1}{2}(\phi_2 \mathbf{n}_2) \times (\phi_1 \mathbf{n}_1) \\ &\quad + \tfrac{1}{12}((\phi_2 \mathbf{n}_2) - (\phi_1 \mathbf{n}_1)) \times ((\phi_2 \mathbf{n}_2) \times (\phi_1 \mathbf{n}_1)) + \cdots.\end{aligned} \tag{44}$$

Note that all we used to establish that $R_2 R_1 = R_3$ for this particular $\mathbf{n}_3$ and ϕ_3 is the exponential form (38) for the rotation matrices and the commutation relations between the t_k matrices. But the $\hat{\mathcal{D}}(\mathbf{n},\phi)$ operators obtain from a similar exponential formula (32),

$$\hat{\mathcal{D}}(\mathbf{n},\phi) = \exp(-i\phi n_k \hat{J}_k) \quad \text{versus} \quad R(\mathbf{n},\phi) = \exp(-i\phi n_k t_k), \tag{45}$$

and the commutation relations between the angular momenta $\hat{J}_k$ are exactly similar to those between the t_k matrices,

$$[\hat{J}_i, \hat{J}_j] = i\epsilon_{ijk}\hat{J}_k \quad \text{versus} \quad [t_i, t_j] = i\epsilon_{ijk} t_k\,. \tag{46}$$

Consequently, repeating the above arguments almost verbatim we get

$$\hat{\mathcal{D}}(\mathbf{n}_2,\phi_2)\hat{\mathcal{D}}(\mathbf{n}_1,\phi_1) = \hat{\mathcal{D}}(\mathbf{n}_3,\phi_3) \tag{47}$$

for exactly the same $\phi_3\mathbf{n}_3$ as in eq. (44). Thus,

$$\textbf{if} \quad R_2 R_1 = R_3 \quad \textbf{then} \quad \hat{\mathcal{D}}(R_2)\hat{\mathcal{D}}(R_1) = \hat{\mathcal{D}}(R_3), \tag{48}$$

which proves that the unitary operators $\hat{\mathcal{D}}(R)$ indeed form a representation of the rotation group.

⋆ ⋆ ⋆

Now consider how various operators of a quantum system transform under rotations. In a picture-independent way, we classify the operators as scalar, vectors, tensors, *etc.*, according to the transformation rules of their matrix elements. Thus, an operator $\hat{S}$ is a scalar operator iff (if and only if) its matrix elements are invariant under rotations,

$$\langle\chi|\,\hat{S}\,|\psi\rangle' \;=\; \langle\chi|\,\hat{\mathcal{D}}^\dagger\hat{S}\hat{\mathcal{D}}\,|\psi\rangle \;=\; \langle\chi|\,\hat{S}\,|\psi\rangle\,. \tag{49}$$

A triplet of operators $\hat{V}_{x,y,z}$ form a vector operator $\hat{\mathbf{V}}$ iff their matrix elements transform into each other as

$$\langle\chi|\,\hat{V}_i\,|\psi\rangle' \;=\; \langle\chi|\,\hat{\mathcal{D}}^\dagger\hat{V}_i\hat{\mathcal{D}}\,|\psi\rangle \;=\; R_{ij}\times\langle\chi|\,\hat{V}_j\,|\psi\rangle\,. \tag{50}$$

Nine operators $\hat{T}_{ij}$ form a 2-index tensor operator iff

$$\langle\chi|\,\hat{T}_{ij}\,|\psi\rangle' \;=\; \langle\chi|\,\hat{\mathcal{D}}^\dagger\hat{T}_{ij}\hat{\mathcal{D}}\,|\psi\rangle \;=\; R_{ik}R_{j\ell}\times\langle\chi|\,\hat{T}_{k\ell}\,|\psi\rangle\,, \tag{51}$$

and likewise for the multi-index tensor operators. And all these formulae must work for any two states $\langle\chi|$ and $|\psi\rangle$, so verifying them for some operator of interest might seem like a challenge.

Fortunately, all these eqs. (49), (50), (51), *etc.*., are equivalent to specific commutation relation of the operator in question with the angular momentum operators which generate the rotation symmetries. Thus (in $\hbar = 1$ units):

$$\begin{aligned}
&\hat{S}\text{ is a scalar operator iff }[\hat{S},\hat{J}_k] \;=\; 0, &(52)\\
&\hat{V}_i\text{ form a vector operator iff }[\hat{V}_j,\hat{J}_k] \;=\; i\epsilon_{jk\ell}\hat{V}_\ell\,, &(53)\\
&\hat{T}_{ij}\text{ form a tensor operator iff }[\hat{T}_{ij},\hat{J}_k] \;=\; i\epsilon_{ik\ell}\hat{T}_{\ell j} \;+\; i\epsilon_{jk\ell}\hat{V}_{i\ell}\,, &(54)\\
&\textit{etc., etc.}
\end{aligned}$$

Now let me prove these criteria for the scalar and the vector operators; to save time and students' patience, I shall skip the tensor case. The scalar case if fairly easy. Indeed, the matrix elements (49) remain invariant for all states $\langle\chi|$ and $|\psi\rangle$ if and only if

$$\hat{\mathcal{D}}^\dagger \hat{S} \hat{\mathcal{D}} \;=\; \hat{S}, \tag{55}$$

which is equivalent to $\hat{S}\hat{\mathcal{D}} = \hat{\mathcal{D}}\hat{S}$ because $\hat{\mathcal{D}}^\dagger\hat{\mathcal{D}} = 1$. On the other hand,

$$\hat{S} \text{ commutes with } \hat{\mathcal{D}} \;=\; \exp(-i\phi n_k \hat{J}_k) \textbf{ for all } \phi\mathbf{n} \tag{56}$$

if and only if $\hat{S}$ commutes with all 3 generators $\hat{J}_k$. Thus, $\hat{S}$ is a scalar operator iff it commutes with the 3 angular momenta, exactly as in the criterion (52).

The vector case is a bit more involved. First, the matrix elements of 3 operators $\hat{V}_i$ transform as vectors according to eq. (50) for all states $\langle\chi|$ and $|\psi\rangle$ iff the operators themselves obey

$$\hat{\mathcal{D}}^\dagger(R)\hat{V}_i\hat{\mathcal{D}}(R) \;=\; R_{ij}\hat{V}_j\,, \tag{57}$$

or in terms of the generators

$$\exp\bigl(+i\phi n_k \hat{J}_k\bigr)\hat{V}_i \exp\bigl(-i\phi n_k \hat{J}_k\bigr) \;=\; R_{ij}(\mathbf{n},\phi)\hat{V}_j\,. \tag{58}$$

Second, by the Campbell identity

$$\begin{aligned} \exp(+A)B\exp(-A) \;&=\; B \;+\; [A,B] \;+\; \tfrac{1}{2}[A,[A,B]] \;+\; \tfrac{1}{6}[A,[A,[A,B]]] \;+\; \cdots \\ &=\; \sum_{N=0}^{\infty}\frac{1}{N!}\,[A,\ldots[A,B]\ldots]_{N\,\text{times}}\,. \end{aligned} \tag{59}$$

for $A = i\phi\mathbf{n}\cdot\hat{\mathbf{J}}$ and $B = \hat{V}_i$, the LHS of eq. (58) becomes

$$\text{LHS} \;=\; \sum_{N=0}^{\infty}\frac{(i\phi)^N}{N!}\,[(\mathbf{n}\cdot\hat{\mathbf{J}}),\ldots[(\mathbf{n}\cdot\hat{\mathbf{J}}),\hat{V}_i]\ldots]_{N\,\text{times}}\,. \tag{60}$$

and $\hat{\mathbf{V}}$ is a vector operator if and only if this LHS $= R_{ij}\hat{V}_j$ for any rotation.

Now consider rotations through infinitesimal angles ϕ. To the first order in ϕ, eq. (60) reduces to

$$\text{LHS} \;=\; \hat{V}_i \;+\; i\phi[(\mathbf{n}\cdot\hat{\mathbf{J}}),\hat{V}_i] \;+\; O(\phi^2) \;=\; \hat{V}_i \;+\; i\phi n_j[\hat{J}_j,\hat{V}_i] \;+\; O(\phi^2). \tag{61}$$

At the same time

$$R \;=\; 1 \;-\; i\phi n_k t_k \;+\; O(\phi^2) \quad\Longrightarrow\quad R_{ij} \;=\; \delta_{ij} \;+\; \phi n_k \epsilon_{ikj} \;+\; O(\phi^2), \tag{62}$$

and therefore

$$\text{RHS} \;=\; R_{ij}\hat{V}_j \;=\; \hat{V}_i \;+\; \phi n_k \epsilon_{ikj}\hat{V}_j \;+\; O(\phi^2). \tag{63}$$

Thus, **if** $\hat{\mathbf{V}}$ is a vector operator **then** we must have LHS = RHS:

$$\hat{V}_i \;+\; i\phi n_j[\hat{J}_j,\hat{V}_i] \;+\; O(\phi^2) \;=\; \hat{V}_i \;+\; \phi n_k \epsilon_{ikj}\hat{V}_j \;+\; O(\phi^2), \tag{64}$$

hence

$$in_j[\hat{J}_j,\hat{V}_i] \;=\; n_k\epsilon_{ikj}\hat{V}_j \;=\; \langle\!\langle\,\text{swapping } j\leftrightarrow k\,\rangle\!\rangle \;=\; n_j\epsilon_{ijk}\hat{V}_k \tag{65}$$

and therefore

$$i[\hat{J}_j,\hat{V}_i] \;=\; \epsilon_{ijk}\hat{V}_k \quad\Longrightarrow\quad [\hat{V}_i,\hat{J}_j] \;=\; i\epsilon_{ijk}\hat{V}_k\,, \tag{66}$$

exactly as in the criterion (53).

The bottom line is, a vector operator must obey the commutation relations (53). To prove the converse, let's go back to finite rotations and the multiple commutator formula

$$\hat{\mathcal{D}}^\dagger\hat{V}_i\hat{\mathcal{D}} \;=\; \sum_{N=0}^{\infty}\frac{(i\phi)^N}{N!}\,[(\mathbf{n}\cdot\hat{\mathbf{J}}),\ldots[(\mathbf{n}\cdot\hat{\mathbf{J}}),\hat{V}_i]\ldots]_{N\,\text{times}}\,. \tag{67}$$

If the operator $\hat{\mathbf{V}}$ obeys the commutation relations (53) then

$$[(\mathbf{n}\cdot\hat{\mathbf{J}}),\hat{V}_i] \;=\; i\epsilon_{ikj}n_j\hat{V}_k \;=\; -(\mathbf{n}\cdot\mathbf{t})_{ik}\hat{V}_k. \tag{68}$$

Consequently,

$$\begin{aligned}[(\mathbf{n}\cdot\hat{\mathbf{J}}),[(\mathbf{n}\cdot\hat{\mathbf{J}}),\hat{V}_k]] &= [(\mathbf{n}\cdot\hat{\mathbf{J}}),-(\mathbf{n}\cdot\mathbf{t})_{ik}\hat{V}_k] = -(\mathbf{n}\cdot\mathbf{t})_{ik}[(\mathbf{n}\cdot\hat{\mathbf{J}}),\hat{V}_k] \\ &= +(\mathbf{n}\cdot\mathbf{t})_{ik}(\mathbf{n}\cdot\mathbf{t})_{k\ell}\hat{V}_\ell = (\mathbf{n}\cdot\mathbf{t})^2_{i\ell}\hat{V}_\ell\,,\end{aligned} \tag{69}$$

and likewise for all multiple commutators:

$$[(\mathbf{n}\cdot\hat{\mathbf{J}}),\ldots[(\mathbf{n}\cdot\hat{\mathbf{J}}),\hat{V}_i]\ldots]_{N\,\text{times}} = (-1)^N(\mathbf{n}\cdot\mathbf{t})^N_{ik}\hat{V}_k\,. \tag{70}$$

Plugging this result into the multiple commutator formula (67) we get

$$\hat{\mathcal{D}}^\dagger\hat{V}_i\hat{\mathcal{D}} = \sum_{N=0}^{\infty}\frac{(-i\phi)^N}{N!}(\mathbf{n}\cdot\mathbf{t})^N_{ik}\hat{V}_k = \left(\sum_{N=0}^{\infty}\frac{(-i\phi)^N}{N!}(\mathbf{n}\cdot\mathbf{t})^N\right)_{ik}\hat{V}_k = \exp(-i\phi(\mathbf{n}\cdot\mathbf{t}))_{ik}\hat{V}_k\,. \tag{71}$$

At the same time, for a finite rotation $R=\exp(-i\phi(\mathbf{n}\cdot\mathbf{t}))$, thus

$$\exp(-i\phi(\mathbf{n}\cdot\mathbf{t}))_{ik}\hat{V}_k = R_{ik}\hat{V}_k \tag{72}$$

and therefore

$$\hat{\mathcal{D}}^\dagger\hat{V}_i\hat{\mathcal{D}} = R_{ik}\hat{V}_k\,. \tag{73}$$

Thus, **if** the $\hat{V}_i$ obey the vector commutation relations (53) with the angular momenta, **then** $\hat{\mathbf{V}}$ transforms under all rotations as a vector operator. *Quod erat demonstrandum.*

Lie Groups, Lie Algebras, and their Representations

A Lie group is a continuous group that's also a differentiable manifold, and both the group product and the inverse must be differentiable functions on the manifold. For example, the rotation group $SO(3)$ has a continuous 3D manifold with coordinates $\phi\mathbf{n}=(\phi n_x,\phi n_y,\phi n_z)$. The group inverse $R^{-1}(\phi\mathbf{n})=R(-\phi\mathbf{n})$ is obviously analytic in these coordinates, and for the group product

$$R((\phi\mathbf{n})_2)\times R((\phi\mathbf{n})_1) = R((\phi\mathbf{n})_3) \tag{74}$$

we also have an analytic formula

$$(\phi\mathbf{n})_3 = (\phi\mathbf{n})_2+(\phi\mathbf{n})_1+\tfrac{1}{2}(\phi\mathbf{n})_2\times(\phi\mathbf{n})_1+\tfrac{1}{12}((\phi\mathbf{n})_2-(\phi\mathbf{n})_1)\times((\phi\mathbf{n})_2\times(\phi\mathbf{n})_1)+\cdots. \tag{75}$$

Consider a general Lie group G of dimension D. Let's parametrize the neighborhood of the unit element 1 by some coordinates $(\alpha_1, \ldots, \alpha_D) = \vec{\alpha}$ which vanish for $g = 1$, so that for infinitesimal $\vec{\alpha}$ we gave

$$g(\vec{\alpha}) \;=\; 1 \;-\; i\sum_{a=1}^{D} \alpha_i t_a \;+\; O(\vec{\alpha}^2) \tag{76}$$

for some *generators*

$$t_a \;=\; i\,\frac{\partial g(\vec{\alpha})}{\partial \alpha_a}\bigg|_{\vec{\alpha}=0}\,. \tag{77}$$

For example, for $G = SO(3)$ we have $D = 3$ and the $t_{1,2,3}$ are the matrices $(t_a)_{ij} = i\epsilon_{iaj}$. For another example, the $SO(N)$ group with $N > 3$ has $D = \frac{1}{2}N(N-1)$ and its generators t_a are the imaginary antisymmetric matrices. The generators themselves are best labeled by antisymmetric pairs $t_{ij} = -t_{ji}$, specifically

$$(t_{ij})_{k\ell} \;=\; i\delta_{ik}\delta_{j\ell} \;-\; i\delta_{i\ell}\delta_{jk}\,, \tag{78}$$

so that for any antisymmetric matrix $A_{ij} = -A_{ji}$ we have

$$\frac{-i}{2}\sum_{ij} A_{ij}t_{ij} \;=\; A. \tag{79}$$

Outside of the infinitesimal neighborhood of 1, we may set the coordinates $(\alpha_1, \ldots, \alpha_D)$ such that the symmetry $g(n\vec{\alpha})$ is the n^{th} power of the $g(\vec{\alpha})$. Consequently, for any finite parameter vector $\vec{\alpha}$

$$\begin{aligned} g(\vec{\alpha}) \;&=\; \lim_{n\to\infty}\big(g(\tfrac{1}{n}\vec{\alpha})\big)^n \\ &=\; \lim_{n\to\infty}\left(1 \,-\, \frac{i\vec{\alpha}\cdot\vec{t}}{n} \,+\, O(1/n^2)\right)^n \\ &=\; \exp(-i\vec{\alpha}\cdot\vec{t}). \end{aligned} \tag{80}$$

We have already seen this formula for the rotation group, and now we see it's rather universal.

Next, consider the group product in this parametrization:

$$g(\vec{\alpha}) \times g(\vec{\beta}) \;=\; g(\vec{\gamma}) \quad \text{for some } \vec{\gamma}, \tag{81}$$

thus

$$\exp(-i\vec{\alpha}\cdot\vec{t}) \times \exp(-i\vec{\beta}\cdot\vec{t}) \;=\; \exp(-i\vec{\gamma}\cdot\vec{t}). \tag{82}$$

By the Baker–Campbell–Hausdorff formula (59), this means

$$\vec{\gamma}\cdot\vec{t} \;=\; (\vec{\alpha}+\vec{\beta})\cdot\vec{t} \;-\; \frac{i}{2}\big[\vec{\alpha}\cdot\vec{t}, \vec{\beta}\cdot\vec{t}\big] \;-\; \frac{1}{12}\big[(\vec{\alpha}-\vec{\beta})\cdot\vec{t}, \big[\vec{\alpha}\cdot\vec{t}, \vec{\beta}\cdot\vec{t}\big]\big] \;+\; \cdots. \tag{83}$$

In light of the LHS here, the RHS must be some linear combination of the generators $(t_1, \ldots, t_D)$; moreover, this must be true for any $\vec{\alpha}$ and $\vec{\beta}$, so treating the RHS as a power series in $\vec{\alpha}$ and $\vec{\beta}$, we find that *every multiple commutator on the RHS of eq.* (83) *must be some linear combination of the generators* $(t_1, \ldots, t_D)$. In particular, the leading commutator

$$\big[\vec{\alpha}\cdot\vec{t}, \vec{\beta}\cdot\vec{t}\big] \;=\; \sum_{a,b} \alpha_a \beta_b \times [t_a, t_b] \;= \text{must be} = \sum_c (\text{coeff})_c t_c \quad \text{for some } (\text{coeff})_c\,. \tag{84}$$

Moreover, this must be true for any $\vec{\alpha}$ and any $\vec{\beta}$, so all the individual commutators of two generators $[t_a, t_b]$ must be linear combinations of some other generators,

$$[t_a, t_b] \;=\; \sum_c i f_{abc} t_c \tag{85}$$

for some coefficients f_{abc} called *the structure constants* of the Lie group. For example, for the $SO(3)$, $f_{abc} = \epsilon_{abc}$.

The commutator algebra (85) of some group's generators is called the *Lie algebra*. In its abstract form, a Lie algebra is defined without any reference to the generators being matrices or operators or to their associative products $t_a t_b$. All one needs are some brackets $[t_a, t_b]$ that act similarly to the commutator brackets: They are antisymmetric and obey the Jacobi identity,

$$[t_a, t_b] \;=\; -[t_b, t_a], \tag{86}$$

$$[t_a, [t_b, t_c]] \;+\; [t_b, [t_c, t_a]] \;+\; [t_c, [t_a, t_b]] \;=\; 0. \tag{87}$$

But for our purposes, you may always think of the t_a being some kind of matrices and the Lie bracket being the commutator bracket.

Given the Lie algebra — that is, its generators and the structure constants — all the multiple commutators in the Baker–Campbell–Hausdorff series (83) obtain as linear combinations of the generators themselves, which gives the power series for the $\vec{\gamma}$ as a function of $\vec{\alpha}$ and $\vec{\beta}$. Thus, the Lie algebra of the group's generators t_a completely determines the group product $g(\vec{\alpha})g(\vec{\beta}) = g(\vec{\gamma})$.

Representations

A representation (r) of a Lie algebra is a set of matrices (of any size) or operators (in any Hilbert space) $T_1^{(r)}, \ldots T_D^{(r)}$ which obey the same commutation relations as the algebra's generators $t_1, \ldots t_D$,

$$\left[T_a^{(r)}, T_b^{(r)}\right] \;=\; i\sum_c f_{abc} T_c^{(r)} \tag{88}$$

for the same structure constants f_{abc} as in eq. (85). Every representation of the Lie algebra is also a representation of the corresponding Lie group,

$$g \;\mapsto\; M^{(r)}(g)$$

where $M^{(r)}$ is a matrix of the same size as the $T_a^{(r)}$ matrices or an operator in the same Hilbert space. Indeed, every $g \in G$ is

$$g(\vec{\alpha}) \;=\; \exp\left(-i\sum\nolimits_a \alpha_a t_a\right) \tag{89}$$

for some parameters $\vec{\alpha} = (\alpha_1, \ldots, \alpha_D)$, so we construct

$$M^{(r)}(g(\vec{\alpha})) \;=\; \exp\left(-i\sum\nolimits_a \alpha_a T_a^{(r)}\right). \tag{90}$$

Since the $T_a^{(r)}$ matrices obey the same commutation relations as the group generators t_a, by the Baker–Campbell–Hausdorff formula we get

$$M^{(r)}(\vec{\alpha})M^{(r)}(\vec{\beta}) \;=\; M^{(r)}(\vec{\gamma}) \tag{91}$$

for exactly the same γ as

$$g(\vec{\alpha})g(\vec{\beta}) \;=\; g(\vec{\gamma}).$$

In terms of the group elements rather than their $\vec{\alpha}$ coordinates, this means

$$M^{(r)}(g_1)M^{(r)}(g_2) \;=\; M^{(r)}(g_1g_2) \quad \forall g_1, g_2 \in G, \tag{92}$$

which makes the $g \mapsto M^{(r)}(g)$ a representation of the Lie group G.

Conversely, any representation (r) of a Lie group is a representation of the corresponding Lie algebra where

$$T_a^{(r)} \;=\; i\frac{dM^{(r)}(g(\vec{\alpha}))}{d\alpha_a}\bigg|_{\vec{\alpha}=0}, \tag{93}$$

so for infinitesimal $\vec{\alpha}$

$$M^{(r)}(g(\vec{\alpha})) \;=\; 1 \;-\; i\sum_a \alpha_a T_a^{(r)} \;+\; O(\alpha^2). \tag{94}$$

Again, the Baker–Campbell–Hausdorff formula assures us that if the matrices $M^{(r)}(g)$ multiply as in eq. (92), then the $T_a^{(r)}$ matrices obey the same commutation relations (88) as the group's generators t_a.

A representation is called *reducible* if all the generators $T_a^{(r)}$ — and hence all the symmetry matrices or operators $M^{(r)}(g)$ — can be simultaneously block-diagonalized (into more than 1 block); if they cannot, the representation is called *irreducible*.⋆ Block diagonalizing all the generators of a reducible representation splits this representation into the tensor sum of smaller representations, one for each diagonal block,

$$(r) \;=\; (r_1) \oplus (r_2) \oplus \cdots \oplus (r_n). \tag{95}$$

⋆ Strictly speaking, the representations with block-diagonalizable generators are called *decomposable* rather than reducible while the representation without block-diagonalization are called *undecomposable* rather than irreducible. The definitions of the reducible and irreducible representations are somewhat different — see the Wikipedia article on the subject, — and some groups may have reducible but undecomposable representations. But for most groups of interest to the Physicists — and in particular all the groups we shall deal with in this class, — all the undecomposable representations are irreducible, all the decomposable representations are reducible, and vice verse. Consequently, most Physicists call the representations reducible or irreducible when they mean decomposable or undecomposable, and I shall follow this inaccurate terminology in this class.

For example, if in some basis for the representation (r)

$$\text{all } T_a^{(r)} \ = \ \begin{pmatrix} T_a^{(r_1)} & 0 \\ 0 & T_a^{(r_2)} \end{pmatrix} \tag{96}$$

then $(r) = (r_1) \oplus (r_2)$. Usually, when block-diagonalizing the generators of a reducible representation one the block that cannot be block-diagonalized any further; in this case, all the component representations $(r_1), (r_2), \ldots, (r_n)$ in the decomposition (95) are irreducible.

For example, the net angular momenta operators $\hat{J}_i = \hat{L}_i + \hat{S}_i$ in the Hilbert space of the hydrogen atom form an infinite-dimensional representation of the $SO(3)$ Lie algebra, hence the $\hat{\mathcal{D}}(\mathbf{n}, \phi)$ rotation operators form the related representation of the $SO(3)$ Lie group. However, this representation is *reducible* and can be block-diagonalized in the $|n, \ell, j, m_j\rangle$ basis, in which all 3 angular momenta $\hat{J}_{x,y,z}$ are block-diagonal WRT to the n, ℓ, and j quantum numbers but not WRT m, thus

$$\hat{J}_i \, |n, \ell, j, m\rangle \ = \ \sum_{m'} \bigl(J_i^{(j)}\bigr)_{mm'} \times \bigl|\text{same}\, n, \ell, j, \ \textit{but not}\, m'\bigr\rangle \,. \tag{97}$$

Consequently, the rotation operators $\hat{\mathcal{D}}(R)$ are block-diagonal in the same basis,

$$\hat{\mathcal{D}}(R) \, |n, \ell, j, m\rangle \ = \ \sum_{m'} \mathcal{D}^{(j)}_{mm'}(R) \times \bigl|\text{same}\, n, \ell, j, \ \text{but not}\, m'\bigr\rangle \,. \tag{98}$$

Multiplets

In Physics terminology, the *multiplets* of some symmetry are often confused with the *representations* of the symmetry group. The two concepts are related but not synonymous, so let me explain the difference.

A *multiplet* is a set of objects — quantum states, operators, fields, whatever — which transform into each other by the symmetry action. For example, consider the $|n, \ell, j, m\rangle$ states of a hydrogen atom. With respect to the $SO(3)$ rotational symmetry group, the states with the same n, ℓ, and j but different m form a multiplet (n, ℓ, j): They transform into each other according to eq. (98), but the transformation law does not involve any states with different n, ℓ, or j. Instead, such states form different multiplets (n',ℓ',j'). In particular, the multiplets (n,ℓ,j) and (n',ℓ',j) — with the same j but different $n'\neq n$ and/or $\ell'\neq\ell$ — are different multiplets comprising different states, even though they transform in a similar way under the $SO(3)$ symmetry.

By comparison, a *representation* is a mathematical object — a homomorphism from the symmetry group G into a space of matrices or linear operators. For example, the homomorphisms

$$R\in SO(3) \quad\mapsto\quad \text{matrix } \hat{\mathcal{D}}^{(j)}_{m,m'}(R) \tag{99}$$

for $j=0,\frac{1}{2},1,\frac{3}{2},2,\ldots$ form finite representations of the $SO(3)$ group, but for any such homomorphism the representation does not care what the $\hat{\mathcal{D}}^{(j)}_{m,m'}(R)$ matrices act upon. Thus, for each j there is only one representation of the $SO(3)$ by $(2j+1)\times(2j+1)$ matrices.

From the multiplet point of view, *a representation is the type of a multiplet*: it specifies how the members of a multiplet transform into each other, regardless of what these members happen to be. In the hydrogen atom example, the multiplets (n,ℓ,j) and (n',ℓ',j) (with the same j but different $n'\neq n$ and/or $\ell'\neq\ell$) are of the same type — the (j) representation of the $SO(3)$ — but they are different multiplets because they comprise different states.

In gauge theory — which we shall study soon — the matter fields form multiplets of the gauge theory group G, and there can be several distinct multiplets of the same type. For example, in QCD (quantum chromodynamics) the gauge group is $SU(3)$ while the matter fields are the quarks — Dirac spinor fields $\Psi^{if}_\alpha(x)$ which come in 3 colors $i=1,2,3$ and 6 flavors $f=u,d,s,c,b,t$. The $SU(3)$ gauge symmetry acts only on the colors,

$$\Psi'^{if}_\alpha(x) \;=\; \sum_j U^i{}_j(x)\Psi^{jf}_\alpha \quad \langle\!\langle\,\text{for the same } f\,\rangle\!\rangle, \tag{100}$$

so the quarks form 6 distinct multiplets according to their flavors: the 3 quark fields of different colors $i=1,2,3$ but of the same flavor form a multiplet. But all 6 of these multiplet are of the same type — the fundamental **3** representation of the $SU(3)$ group.