Symmetries, Redundancies, and Superselection

Domenico Giulini
Institute of Physics
University of Freiburg

American Institute of Physics, College Park, April 30.

Abstract

In physics the notion of "symmetry" comes in various guises, which differ significantly in meaning as well as in their physical and mathematical consequences. It is a potentially confusing fact that these different notions may coexist within a single group of "symmetries", as it happens in the case of gauge or generally covariant theories (Yang-Mills, General Relativity). I will explain how certain statements concerning the existence of superselection rules seem to rest on a conflation of these notions, thereby "degrading" them to the level of ordinary selection rules. This may be welcomed on physical grounds. Proofs of superselection rules using projective representations are also considered.

Equations of Motion

- Dynamical laws usually contain two types of objects:
  1. Background structures, Σ, which are externally specified, and
  2. dynamical structures, Φ, like 'particles' and 'fields', which are solved for.
- Equations of motion then establish a relation between these structures:

$$\mathrm{EM}[\Phi, \Sigma] = 0\,. \tag{1}$$

- If $\mathcal{K}$ is the space of all Φ considered ("kinematically possible trajectories"), then eq. (1) is read as selecting a subset $\mathcal{D} \subset \mathcal{K}$ ("dynamically possible trajectories") for *given* Σ.

Symmetries

- Suppose a group G acts on $\mathcal{K}$:

$$G \times \mathcal{K} \to \mathcal{K}\,, \quad (g, \Phi) \mapsto g \cdot \Phi\,. \tag{2}$$

We call G a group of *symmetries* if it leaves $\mathcal{D} \subset \mathcal{K}$ invariant (as a subset). In other words, if for all $g \in G$:

$$EM[\Phi\,,\,\Sigma] = 0 \iff E[g \cdot \Phi\,,\,\Sigma] = 0\,. \tag{3}$$

- This is to be distinguished from mere 'covariance', which just states the invariance of the relation established by EM:

$$EM[\Phi\,,\,\Sigma] = 0 \iff E[g \cdot \Phi\,,\,g \cdot \Sigma] = 0\,. \tag{4}$$

Proper Symmetries and Redundancies

- If, for each $g \in G - \{1\}$, there exists a $\Phi \in \mathcal{D}$ such that $g \cdot \Phi$ is physically distinguishable from Φ, we would call G a group of *proper* physical symmetries.
- If, on the other hand, $g \cdot \Phi$ is physically indistinguishable from Φ for all $g \in G$ and all $\Phi \in \mathcal{D}$, then we would call G a group of *gauge transformations* or *redundancies*.
- Typical situation encountered in physics are mixtures of these two extreme cases: the equations of motion allow for a group G of symmetries which contains a normal subgroup, Gau, of gauge transformations. Transformations in $G - \mathrm{Gau}$ correspond to proper physical symmetries. The quotient group $\mathcal{S}\mathrm{ym} := G/\mathrm{Gau}$ will then act by proper physical symmetries on reduced phase space. Mathematically expressed, the three groups in question form a short exact sequence

$$1 \longrightarrow \mathrm{Gau} \stackrel{i}{\longrightarrow} G \stackrel{q}{\longrightarrow} \mathcal{S}\mathrm{ym} \longrightarrow 1 \qquad (5)$$

- However, real life (of a theorist) is often more subtle, since the precise size of Gau (and hence of $\mathcal{S}\mathrm{ym}$) may depend on interpretational issues.

Who decides on Gau?

- The existence of gauge symmetries is revealed in the formalism by equations of motion which are under-determining. In the (orthodox) Dirac-Hamilton formalism there will be a bijection between gauge symmetries and (first-class) constraints.
- If consistently carried through, the Hamilton formalism should give an unambiguous answer to the question of what Gau is: it is the group generated by the constraints—and no more! However, in field theory a locality principle (added 'by hand') effectively further constrains the observable content of the theory. It is pretended that Gau is larger than defined above (e.g. 'rigid' gauge transformations.) But this is, though not inconsistent, at least not 'in the spirit' of the equations of motion.
- The existence of superselection rules, like that for electric charge, is based on such *additional* restrictions of observables. They do *not* follow from the constraints of the Dirac-Hamilton formalism.

The Charge SSR I

- The charge SSR states that the charge operator, $\hat{Q}$, lies in the centre of the algebra $\mathcal{A}_{\rm phys}$ of physical observables: $[\hat{Q}, \hat{A}] = 0$ for all $\hat{A} \in \mathcal{A}_{\rm phys}$.
- As a result, states for different charge, Ψ_q and $\Psi_{q'}$, where $\hat{Q}\Psi_q = q\Psi_q$ and $\hat{Q}\Psi_{q'} = q'\Psi_{q'}$ with $q \neq q'$, are *disjoint* relative to $\mathcal{A}_{\rm phys}$ ("fall into different sectors"):

$$0 = \langle \Psi_q \mid [\hat{Q}, \hat{A}] \mid \Psi_{q'} \rangle = (q - q')\langle \Psi_q \mid \hat{A} \mid \Psi_{q'} \rangle \Leftrightarrow \langle \Psi_q \mid \hat{A} \mid \Psi_{q'} \rangle = 0 \,. \tag{6}$$

- Alternatively one may say that the mathematical superposition $\Psi := (\Psi_q + \Psi_{q'})/\sqrt{2}$ is a *mixed* state for $\mathcal{A}_{\rm phys}$. Indeed, for all $\hat{A} \in \mathcal{A}_{\rm phys}$ we have

$$\langle \Psi \mid \hat{A} \mid \Psi \rangle = \mathrm{Tr}(\hat{\rho}\hat{A}) \,, \quad \text{where} \quad \hat{\rho} = \tfrac{1}{2}\big(\mid \Psi_q \rangle\langle \Psi_q \mid + \mid \Psi_{q'} \rangle\langle \Psi_{q'} \mid\big) \,, \tag{7}$$

showing that for $\mathcal{A}_{\rm phys}$ the vector $|\Psi\rangle$ defines a state that is a non-trivial convex combination of other states and hence mixed.

The Charge SSR II

- A proof of the SSR for electric charge was given in 1974 by Strocchi and Wightman in the framework of local relativistic field theory (Haag–Kastler theory). Its basic physical idea rests on Gauß' law, which implies that the total charge may be measured by the flux through *arbitrarily large* spheres. Have

$$\hat{Q} := \lim_{R\to\infty} \int_{\|\mathbf{x}\|\leq R} \hat{\rho}(\mathbf{x}, t)\, d^3x \qquad \text{and} \qquad \nabla \cdot \hat{\mathbf{E}} = \hat{\rho}\,, \tag{8}$$

so that

$$[\,\hat{Q}, \hat{A}\,] = \lim_{R\to\infty} \int_{\|\mathbf{x}\|\leq R} [\,\hat{\rho}(\mathbf{x}, t)\,,\, \hat{A}\,]\, d^3x = \lim_{R\to\infty} \int_{\|\mathbf{x}\|=R} [\,\mathbf{n}\cdot\hat{\mathbf{E}}(\mathbf{x}, t)\,,\, \hat{A}\,]\, d\sigma\,. \tag{9}$$

- Locality of $\mathcal{A}_{\rm phys}$ now implies $[\hat{Q}, \hat{A}] = 0$, since the two-sphere $\|\mathbf{x}\| = R$ is eventually pushed outside the causal complement of the (bounded) support of $\hat{A}$.
- The technical difficulty consists in showing that the formulae (8) and (9) make sense if things have hats on.

Taking the Action Principle Seriously I

- Consider electromagnetism in the spatially bounded region $B_R := \{\|\mathbf{x}\| \leq R\}$. A suitable Lagrangian, *which allows for stationary points even if the normal component of the electric field does not vanish on the spatial boundary* $S_R := \partial B_R$, is as follows (important in QED; Gervais & Zwanziger 1980):

$$
\begin{aligned}
L = &\int_{B_R} \left\{ \dot{\mathbf{A}} \cdot (-\mathbf{E}) - \left[\tfrac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2) + \phi(\rho - \nabla \cdot \mathbf{E}) - \mathbf{A} \cdot \mathbf{j} \right] \right\} d^3x \\
&+ \int_{S_R} \left\{ \dot{\lambda}\, f - \phi \left(R^2 \mathbf{n} \cdot \mathbf{E} - f \right) \right\} \sin\theta \, d\theta d\varphi
\end{aligned}
\tag{10}
$$

- The pair (λ, f) of canonical coordinates label field degrees of freedom on S_R. Their introduction allows to maintain gauge invariance up to, and including, the boundary S_R.
- The field ϕ is a Lagrange multiplier that generates gauge transformations.
- The surface term is necessary to make the action functional differentiable with respect to $\mathbf{E}$ in the class of functions in which $\mathbf{n} \cdot \mathbf{E} \neq 0$ on S_R.
- *This necessity remains in the limit $R \to \infty$ if globally charged configurations are to be included in the domain of differentiability for the action.*

Taking the Action Principle Seriously II

- Expressing the surface fields (λ, f) in terms of their components (λ_{lm}, f_{lm}) with respect to spherical harmonics Y_{lm}, and also writing E_{lm} and ϕ_{lm} for the components of $(R^2\mathbf{n}\cdot\mathbf{E})$ and ϕ respectively, the Hamiltonian reads

$$H = \int_{B_R} \left\{ \tfrac{1}{2}(\mathbf{E}^2 + \mathbf{B}^2) + \phi(\rho - \nabla\cdot\mathbf{E}) - \mathbf{A}\cdot\mathbf{j} \right\} d^3x + \sum_{lm} \phi_{lm}(E_{lm} - f_{lm})\,. \tag{11}$$

- The equations of motion that follow from it are:

$$\dot{\mathbf{A}} = \frac{\delta H}{\delta(-\mathbf{E})} = -\mathbf{E} - \nabla\phi\,, \qquad -\dot{\mathbf{E}} = -\frac{\delta H}{\delta \mathbf{A}} = \mathbf{j} - \nabla\times\mathbf{B}\,, \tag{12}$$

$$\dot{\lambda}_{lm} = \frac{\partial H}{\partial f_{lm}} = -\phi_{lm}\,, \qquad \dot{f}_{lm} = -\frac{\partial H}{\partial \lambda_{lm}} = 0\,. \tag{13}$$

- Variation with respect to ϕ (bulk) and ϕ_{lm} (boundary) leads to Gauß constraints

$$G(\mathbf{x}) := \nabla\cdot\mathbf{E}(\mathbf{x}) - \rho(\mathbf{x}) = 0\,, \qquad G_{lm} := E_{lm} - f_{lm} = 0\,. \tag{14}$$

Formal Schrödinger Quantisation

- We consider a Schrödinger state functional, $\Psi\big(\mathbf{A}(\mathbf{x}), \lambda_{lm}\big)$ over the space of configurations.
- The momentum operators become derivatives in the standard (Dirac) fashion

$$-\mathbf{E}(\mathbf{x}) \to -\hat{\mathbf{E}}(\mathbf{x}) =: -i\hbar \frac{\delta}{\delta \mathbf{A}(\mathbf{x})}, \qquad f_{lm} \to \hat{f}_{lm} := -i\hbar \frac{\partial}{\partial \lambda_{lm}}. \tag{15}$$

- The quantised constraints lead to the standard Gauß constraint in the bulk and

$$\hat{E}_{lm}\Psi = -i\hbar\, \partial\Psi/\partial\lambda_{lm}. \tag{16}$$

- For $l = m = 0$ we have $\hat{E}_{00} = \hat{Q}/\sqrt{4\pi}$; hence

$$\hat{Q}\Psi = \big(-i\hbar\sqrt{4\pi}\big)\, \partial\psi/\partial\lambda_{00}. \tag{17}$$

The Charge SSR Revisited

- A charge SSR, stating that all physical observables commute with the charge operator, is equivalent to the physical impossibility to localise the system in the λ_{00} coordinate.
- Generally, the restriction to (quasi-) local observables forces all $\hat{f}_{lm}$ into the centre of the algebra of observables, that is, it effectively removes the $\hat{\lambda}_{lm}$ from the algebra of observables. This rich superselection structure is precisely what is obtained in more rigorous approaches to local QED (e.g. Buchholz 1982).
- However, from the consistent-action-principle point of view there is no ambiguity as to the 'real' nature of the degrees of freedom labelled by λ_{lm}. For example, a real-time motion in λ_{00} (generated by Q) produces a non-zero action (by definition of $Q \neq 0$). *It is therefore a symmetry, not a redundancy!*
- If the λ_{lm} are effectively unmeasurable, then this must be a *contingent dynamical limitation*. Mere kinematical gauge-invariance cannot be held responsible for that. For example, localisation in the λ_{lm} degrees of freedom may well be unstable against environmental decoherence (see Joos *et al.* 2003).

Analogies in GR

- The gravitational field of (approximately) isolated systems is described by spatially asymptotically Minkowskian space-time geometries.
- They possess global conserved quantities which may be expressed by spatial surface integrals (canonical variables (g_{ab}, π^{ab})):

$$\langle P, \xi \rangle = \lim_{R \to \infty} \int_{S_R} d\omega R^2 \pi(n, \xi) \,, \qquad E = \lim_{R \to \infty} \int_{S_R} d\omega R^2 n^a \big(\partial_b g_{ab} - \partial_a g_{bb} \big) \,. \quad (18)$$

- Are we to postulate SSRs for these Poincaré charges: linear momentum, angular momentum, mass-energy, ...) ? Certainly not, since we think of the conjugate variables (the λ's) as coordinates with respect to a physically realised asymptotic reference frame, measuring position, orientation and time, e.g. by 'looking at the fixed stars'. This leaves no interpretational doubt that the Poincaré charges generate proper physical symmetries.
- In analogy, the variable λ in QED would have to be interpreted as a *relative* phase (Aharonov & Susskind 1967, Mirman 1969, ...). Compare discussion of univ.–SSR.

Topological Issues

- Consider generally Yang-Mills and/or diffeomorphism-invariant theories, describing isolated systems with global charges; hence G is the group of gauge transformations and/or diffeomorphisms. The group generated by the constraints is $\mathcal{G}au = G^0_\infty$, the identity component of all asymptotically trivial symmetry transformations.
- If we stick to the philosophy that redundancies are precisely those transformations generated by the constraints, the group of proper physical symmetries is given by:

$$\mathcal{S}ym := G/G^0_\infty = \left(G/G_\infty\right) \,\tilde{\times}\, \left(G_\infty/G^0_\infty\right) , \tag{19}$$

combining a continuous part, $\mathcal{S}_c := G/G_\infty$, with a discrete part, $\mathcal{S}_d := G_\infty/G^0_\infty$.

- This combination of $\mathcal{S}_c$ with $\mathcal{S}_d$ might be non-trivial, i.e. not just a direct product $\times$; hence we wrote $\tilde{\times}$. Technically speaking, $\mathcal{S}ym$ is an extension of $\mathcal{S}_c$ by $\mathcal{S}_d$. What extension depends on topological issues – with interesting consequences.

Interesting Consequences

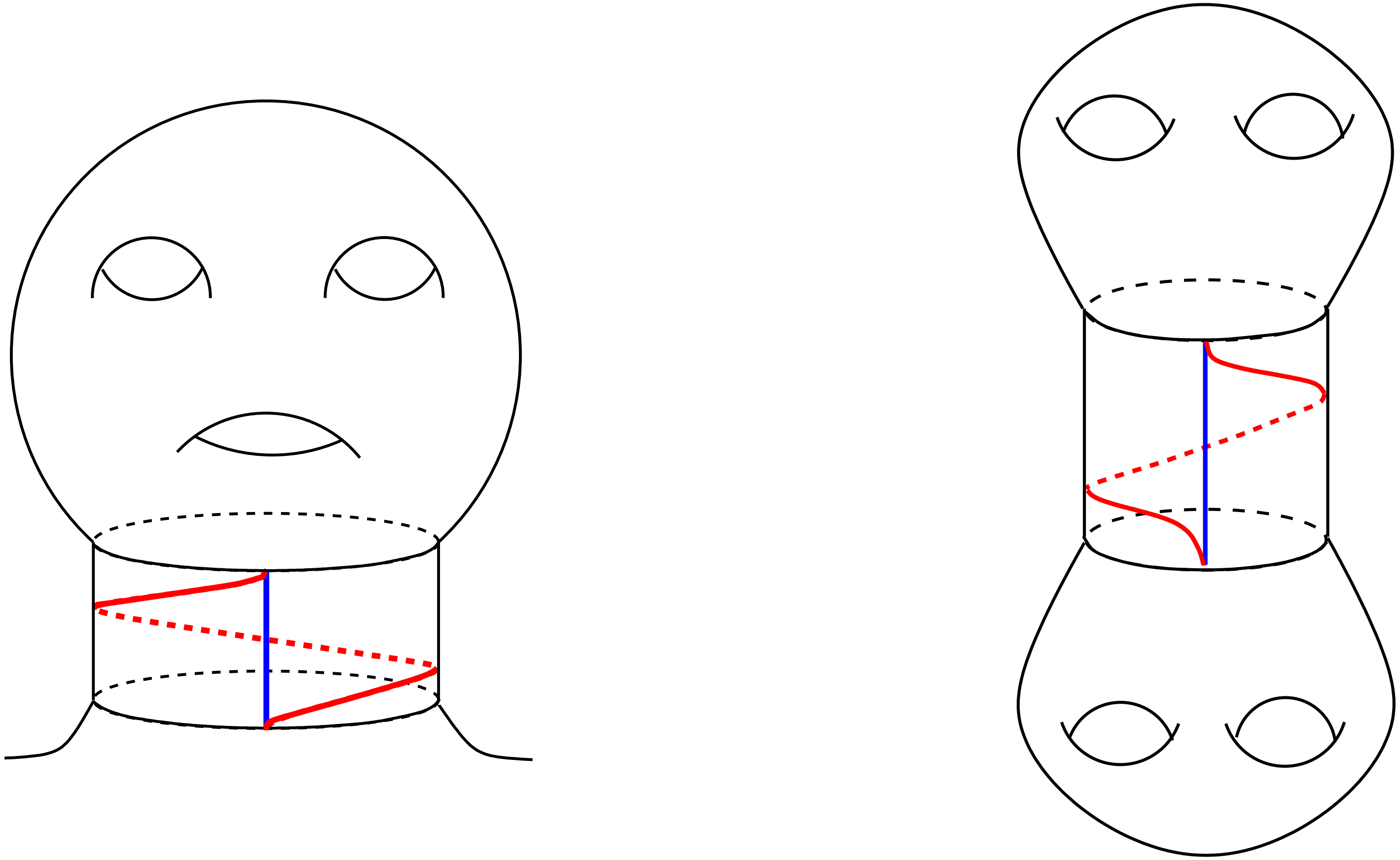
- The $SU(2)$-Yang-Mills-Higgs system allows for regular monopole solutions with both, electric and magnetic charge—so called *dyons* (Julia & Zee 1975). Here, $S_c \cong U(1)$, $S_d \cong \mathbb{Z}$ and (D.G. 1995)

$$\mathcal{Sym} \cong \mathbb{R} \times \mathbb{Z}_{|m|} \,. \tag{20}$$

  The discrete part renders the continuous part non-compact, hence giving rise to the possibility of fractional charge quantisation (Witten 1979).

- In GR, asymptotically flat initial data surfaces may be realised with a wide variety of topologies. Here S_c contains the group of spatial rotations, $SO(3)$, but $\mathcal{Sym}$ may only contain its double cover, $SU(2)$, depending on the topology of the Cauchy surface. This opens up the possibility to have half-integer angular-momentum states in pure gravity (Friedman & Sorkin 1980).

Spinorial 3-Manifolds



Projective Representations

- Symmetries are often realised by projective unitary representations:

$$U(g_1)U(g_2) = \omega(g_1, g_2)\, U(g_1 g_2)\,, \tag{21}$$

where $\omega : G \times G \to U(1)$ cannot be removed by redefinitions $U(g) \mapsto U'(g) := \phi(g)U(g)$ with $\phi : G \to U(1)$.

- This is the same as saying that rather than G, a (central) $U(1)$ extension G' of it acts by proper representations. As set $G = U(1) \times G$, with multiplication given by

$$\big(a_1, g_1\big)\big(a_1, g_2\big) = \big(a_1 a_2 \omega(g_1, g_2)\,,\; g_1 g_2\big)\,. \tag{22}$$

Projective Representations and SSR

- **Theorem:** Two projective unitary representations U', U'' can be subrepresentations of a (reducible) projective representation U if and only if their multipliers are similar.
- For the proof, just consider

$$U(g_1)U(g_2) = \omega'(g_1, g_2)\, U'(g_1 g_2) \oplus \omega''(g_1, g_2)\, U''(g_1 g_2)\,. \tag{23}$$

Univalence SSR

- This applies to $SO(3)$ invariance in QM, where it gives rise to the *univalence* SSR. A proof is as follows: consider spin 1/2 representation, restricted to subgroup of 180° rotations C_i about x, y, z axes, which is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. Then have projective representation

$$C_i \to \exp\big(-i\pi\sigma_i/2\big)\,, \quad \text{with} \quad \omega(C_i, C_i) = \omega(C_i, C_j) = -\omega(C_j, C_i) = 1\,. \qquad (24)$$

- Clearly, no redefinition can remove the ω,

$$\omega'(C_i, C_j) = \omega(C_i, C_j)\,\phi(C_iC_j)/\big[\phi(C_i)\phi(C_j)\big]\,, \qquad (25)$$

  since ω is antisymmetric in C_i, C_j ($i \neq j$), whereas the factor involving ϕ's is symmetric.

Galilean Symmetry

- A similar Argument applies to Galilean symmetry in non-relativistic QM, e.g. for of n particles with Galilei invariant potential.
- A typical group element is $g = (R, \mathbf{v}, \mathbf{a}, b)$, and the multipliers are given by

$$\omega(g_1, g_2) = \exp\left\{\tfrac{i}{\hbar} M\, \xi(g_1, g_2)\right\}\,, \qquad \xi(g_1, g_2) = \mathbf{v}' \cdot R' \cdot \mathbf{a} - \tfrac{1}{2}\mathbf{v}'^2 b\,, \tag{26}$$

  where $M = \sum_i m_i$ is the total mass. Hence get a SSR for states with different total mass, as stated in some good textbooks on QM.
- However, does that statement actually make sense if mass is not a dynamical variable? What is the system two states of which might correspond to different overall mass?

Appearence of Schrödinger Group

- Add a canonical pair of variables (λ_i, m_i), $i = 1, \cdots, n$, for each point mass and consider minimally extended action:

$$S = \int dt \left(\sum_{i=1}^{n} m_i \dot{\lambda}_i + \sum_{i=1}^{n} \mathbf{p}_i \cdot \dot{\mathbf{x}}_i \; - \; H\big(\{\mathbf{x}_a\}, \{\mathbf{p}_a\}, \{m_a\}\big) \right) \tag{27}$$

- The equations of motions for $\big(\{\mathbf{x}_i\}, \{\mathbf{p}_i\}\big)$ are as before, whereas the ones for the new variables read:

$$\dot{\lambda}_i = \frac{\partial V}{\partial m_i} - \frac{\mathbf{p}_i^2}{2m_i^2}, \qquad \dot{m}_i = 0\,. \tag{28}$$

- The canonical symmetry group of that dynamical system is not Gal, but rather its universal central extension, Schrö (Schrödinger group), which does *not* give rise to any SSR:

$$1 \rightarrow \mathbb{R} \rightarrow \text{Schrö} \rightarrow Gal \rightarrow 1 \tag{29}$$

Conclusion

- Many superselection rules do not seem to be “super” after all.
- Their apparent existence should be understood in terms of contingent dynamical processes, like e.g. decoherence.
- Proofs of SSR on the basis of projective representations depend on ones classical prejudice of what the “right” symmetry group is.
- Existence of SSR directly relaties to locality principles and absolute-versus-relative issues (Wigner, Hegerfeld,.. $\leftrightarrow$ Aharonov, Susskind, Lubkin, Mirman,...)

END