AMSI VACATION RESEARCH SCHOLARSHIPS 2018-2019

The Geometric and Probabilistic Structure of Classical Physical Theories

Shay Tobin

Supervised by Frank Valckenborgh

Macquarie University

28/02/2019

Vacation Research Scholarships are funded jointly by the Department of Education and Training and the Australian Mathematical Sciences Institute.

AMSI AUSTRALIAN MATHEMATICAL SCIENCES INSTITUTE

Australian Government
Department of Education and Training

1 Introduction

Physical systems can be characterised operationally by how they interact with measurement devices associated with experiments on the system and the mathematical models used to depict such systems should reflect this interaction. Classical probability theory can be utilised for the description of the events and their likelihoods, associated with all possible outcomes for a *single* experiment, as a probability space $(\mathcal{E}, \mathbb{P})$, where $\mathbb{P}$ is a probability measure on $\mathcal{E}$, the collection of all possible events; this is always a σ-complete distributive complemented lattice. A suitable mathematical model then for a physical system has to integrate all information from all possible experiments on that system in a single mathematical construct in a consistent, optimal and preferably essentially unique way. The general mathematical description of a physical system introduces the concepts of a state space Σ (states assigned to the system) and the collection of observables $\mathcal{O}$ that represent the measurements, and this leads to the so-called propositional calculus $\mathcal{L}$ (Birkhoff & von Neumann, 1936; Jauch, 1968).

On the other hand, the familiar mathematical model for a classical physical system carries much more geometric structure. One usually starts with constructing so-called *configuration spaces*, *phase spaces* and *state spaces*. These are smooth manifolds with additional geometric structures (Kibble & Berkshire, 2004; Loomis & Sternberg, 1968). Properties assigned to the system under investigation correspond with (measurable) subsets of these spaces. Observables are essentially functions on phase space or state space (the set of possible states of the system). In particular, the Hamiltonian function generates the dynamics of the system. Alternatively, the Lagrangian approach can also be used to generate the dynamical equations.

The motivation for this project is to construct parts of the bridge between the general probabilistic framework, starting from the general pair of states and properties or propositions $(\Sigma, \mathcal{L})$, and the more familiar geometric structures used for the description of classical physical systems. In fact, I will start with making these ideas precise for the concrete example of an elementary point particle moving in $\mathbb{R}^3$, one of the simplest non-trivial classical systems, and investigate the essential uniqueness of the usual mathematical model of this system, consisting of its *configuration space* $U \subseteq \mathbb{R}^3$ (induced by the position observable) and the associated tangent and cotangent bundles; the *phase space* T^*U corresponds with the cotangent bundle of U and the Hamiltonian lives at this level. In addition, one can construct its *state space* $T^*U \times \mathbb{R}$.

The dynamics of the system is generated by the Hamiltonian function defined on phase space of the general form

$$H(\mathbf{q},\mathbf{p}) = \frac{(\mathbf{p}-\mathbf{A}(\mathbf{q}))^2}{2m} + V(\mathbf{q})$$

or on state space with

$$H(\mathbf{q},\mathbf{p},t) = \frac{(\mathbf{p}-\mathbf{A}(\mathbf{q},t))^2}{2m} + V(\mathbf{q},t)\,.$$

These expressions can be obtained from imposing the Galilean requirement that changes in momentum $\mathbf{p}\mapsto\mathbf{p}+\mathbf{w}$ are associated with changes in velocity of the form $\dot{\mathbf{q}}\mapsto\dot{\mathbf{q}}+\frac{1}{m}\mathbf{w}$ (Piron, 1998). Explicitly, it then follows from Hamilton's equations on phase space

$$\begin{aligned}\frac{\mathrm{d}q^i}{\mathrm{d}t}(\mathbf{q},\mathbf{p}) &= \frac{\partial H}{\partial p_i}(\mathbf{q},\mathbf{p})\\ \frac{\mathrm{d}p_i}{\mathrm{d}t}(\mathbf{q},\mathbf{p}) &= -\frac{\partial H}{\partial q^i}(\mathbf{q},\mathbf{p})\end{aligned}$$

that

$$\frac{\partial H}{\partial p_i}(\mathbf{q},\mathbf{p}+\mathbf{w}) = \frac{\partial H}{\partial p_i}(\mathbf{q},\mathbf{p}) + \tfrac{1}{m}w_i\,.$$

This is a functional equation of the form $f(x+w) = f(x)+\frac{1}{m}w$ valid for all x and w, with general solution $f(y) = \frac{1}{m}y + f(0)$, and so we have in our case

$$\frac{\partial H}{\partial p_i}(\mathbf{q},\mathbf{p}) = \frac{1}{m}\left(p_i - A_i(\mathbf{q})\right)$$

for a suitable function $\mathbf{A}(\mathbf{q})$ from which the Hamiltonian above follows by a simple integration. A similar argument establishes the general form of the Hamiltonian on state spaces.

Specifically, we will start with the observation that the elementary point particle can be completely characterised by the two observables of *position* and *linear momentum*, which associate properties of the corresponding measurement devices encoded in the Borel sets $\mathcal{B}(\mathbb{R}^3)$ with properties of the physical system as encoded in $(\Sigma,\mathcal{L})$. These observables must satisfy additional symmetry constraints imposed by relativity considerations, mathematically encoded by so-called systems of imprimitivity (Foulis & Wilce, 2000). For each observable, we have a commutative diagram of the form

$$\begin{array}{ccc}\mathcal{B}(\mathbb{R}^3) & \xrightarrow{\ \mu\ } & (\Sigma,\mathcal{L})\\ {\scriptstyle g}\downarrow & & \downarrow{\scriptstyle g}\\ \mathcal{B}(\mathbb{R}^3) & \xrightarrow[\ \mu\ ]{} & (\Sigma,\mathcal{L})\end{array} \tag{1.1}$$

Here, μ denotes a given observable (a morphism of σ-algebras), and g denotes the action $\mathbf{x} \mapsto g \cdot \mathbf{x}$ of an appropriate symmetry group G at the levels of the events associated with the measurement device characterising the observable μ on the one hand, and the mathematical model of the physical system on the other. In our case, G will be a classical symmetry group, the Newton group (also referred to as the passive Galilei group), and the important observables will be the position and momentum observables.

2 The general description of a physical system

In this section, we follow Piron (1998) and Moore (1999). According to this approach, a physical system is characterised by how it interacts with measurement devices and this is formalised using the primitive notions of property, pure state and observable.
A physical system $\mathcal{G}$ can be measured using appropriate measuring devices. Each experimental setting is associated with an experimental project α from a collection of possible experiments $\mathscr{Q}$. A *definite experimental project* relative to a physical system is a real experimental procedure where a positive response from the experiment has been defined in advance and should the experiment be performed: satisfying these conditions yields the response 'yes', if these conditions are not met then the response is 'no'. Restricting the possible questions asked to *(yes/no) questions* is reasonable since one can analyse an experimental procedure with a well defined set Υ of possible responses by considering the definite experimental project α_A for $A \subseteq \Upsilon$, where we assign the response 'yes' if the result is in A (Moore, 1999).

A definite experimental project α is said to be *certain* for a given realisation of $\mathcal{G}$ if when performed always gives a yes response. Consequently one can expect with certainty for α to give a yes response regardless if the experiment is performed or not. This counterfactual way of defining *certainty* allows one to assign multiple properties to the system, even if some of those properties are incompatible. We will use the notation $p \triangleleft \alpha$ if α is certain for $\mathcal{G}$ in state p.

A natural physical assumption that relates to the duality between states and properties is that the collection of all experiments allows one to distinguish between distinct pure states, in other words the relation $\triangleleft$ must satisfy the following property:

$$\text{If } p, q \in \Sigma \text{ and } p \neq q\,, \text{ there exists } \alpha \in \mathscr{Q} \text{ with } p \triangleleft \alpha \text{ and } q \not\triangleleft \alpha. \tag{2.1}$$

The notions of pure states and properties as well as the duality between them is explained in the following.

Let $\alpha, \beta \in \mathscr{Q}$ where $\mathscr{Q}$ is the collection of all possible experimental projects that can be performed on a given preperation of $\mathcal{G}$. A relation $\prec$ is written $\alpha \prec \beta$ and says β is certain in each case α is certain. In terms of the previous notation, we have $\alpha \prec \beta$ if and only if $p \triangleleft \alpha$ implies $p \triangleleft \beta$. The relation $\prec$ is trivially a pre-order on $\mathscr{Q}$. The definite experimental project I is the maximal element of $\mathscr{Q}$ and is defined to be *the experimental project that assigns the response "yes" regardless of what you do to the system* while the minimal element is provided by the definite experimental project O and is defined to be *the experimental project that assign the response "no" regardless of what you do to the system.* An equivalence relation $\approx$ is then established

$$\alpha \approx \beta \text{ if } \alpha \prec \beta \text{ and } \beta \prec \alpha. \tag{2.2}$$

A *product* Π of a family of definite experimental projects $\mathbf{A} \in \mathscr{Q}$ is written $\Pi\mathbf{A}$ and is said to be certain whenever all $\alpha \in \mathbf{A}$ are certain.
An *inverse* $\alpha^{\sim}$ is equivalent to switching the yes/no responses on α such that

$$(\alpha^{\sim})^{\sim} = \alpha \tag{2.3}$$

$$(\Pi\mathbf{A})^{\sim} = \Pi(\mathbf{A}^{\sim}). \tag{2.4}$$

It is important to note that the inverse project $\alpha^{\sim}$ attributes no logical meaning with regards to certainty since the statement that α is not certain does not assume any relation to the certainty of $\alpha^{\sim}$. This is made explicit using the following example. Given the two definite experimental projects O and $O\Pi I$, neither of these is ever certain but $O^{\sim} = I$ which is always certain where as $(O\Pi I)^{\sim} = I\Pi O$ which continues to never be certain. That is, the unary relation $\sim$ on $\mathscr{Q}$ does not descend to the quotient.

For any physical system, there will be associated to it, particular definite experimental projects which are certain for the same realisations. This leads to the notion of property.

Definition 1. A *property* $[\alpha]$ associated with the definite experimental project α is the set of all definite experimental projects β such that $\alpha \approx \beta$.

$$[\alpha] = \{\beta \in \mathscr{Q} \mid \alpha \approx \beta\} \tag{2.5}$$

Proposition 1. *The relation $\prec$ is a preorder and induces a partial order $\leqslant$ on the collection of properties $\mathscr{L}$. The product operation Π equips $\mathscr{L}$ with enough mathematical structure to be a complete lattice.*

The following is a brief overview of the reasoning given in Moore (1999) to the above statement. For $\mathscr{L}$ to be a complete lattice it must have a greatest lower bound $\bigwedge$. This is produced by the product operation Π such that

$$\bigwedge\{[\alpha] \mid \alpha \in \mathbf{A}\} = [\Pi\mathbf{A}]. \tag{2.6}$$

$\mathscr{L}$ must also have a least upper bound $\bigvee$ this is automatically given by Birkhoff's theorem which states that each complete meet semi-lattice is a complete lattice.

$$\bigvee \mathscr{A} = \bigwedge\{b \in \mathscr{L} \mid (\forall a \in \mathscr{A}) a \leqslant b\} \tag{2.7}$$

A pure state of a physical system can be thought of as the singular realisation of a particular physical system. The pure states p and q from the collection of pure states Σ are *orthogonal* $\perp$ if and only if there exists $\alpha \in \mathscr{Q}$ such that $p \triangleleft \alpha$ and $q \triangleleft \alpha^{\sim}$.

Due to the implemented duality between pure states and properties as formulated previously, the orthogonality between states induces a orthogonality between properties. For $a, b \in \mathscr{L}$

$$a \perp b \text{ if and only if } p \perp q \text{ for all } p \triangleleft a \text{ and } q \triangleleft b. \tag{2.8}$$

General states can be defined as follows for the physical system.

Definition 2. A *state* on a property lattice is a generalised probability measure μ on $\mathscr{L}$. Explicitly, we require

(1) $\mu(I) = 1$,

(2) If $i \mapsto a_i$ is a family of pairwise orthogonal properties, then

$$\mu\left(\bigvee_i a_i\right) = \sum_i \mu(a_i), \tag{2.9}$$

(3) If $\mu(a) = 1$ and $\mu(b) = 1$, then $\mu(a \wedge b) = 1$.

The last requirement reflects the physical interpretation of the meet in property lattices. Observe that a state reduces to a classical probability measure when restricted to Boolean σ-algebras of $\mathscr{L}$. From this perspective, a state p is pure if there exists an atom $a_p \in \mathscr{L}$ such that $p(a_p) = 1$.

The duality between the notions of property and pure state is formalised in the *Cartan map*

$$\mu : \mathscr{L} \to \mathcal{P}(\Sigma) : a \mapsto \{p \in \Sigma \mid p \triangleleft a\} . \tag{2.10}$$

It turns out that the Cartan map is injective, order-preserving and $\mu\left(\bigwedge_i a_i\right) = \bigcap_i \mu(a_i)$.[1]
A *question* $\gamma \in \mathscr{Q}$ is said to be *classical* if either $p \triangleleft \gamma$ or $p \triangleleft \gamma^{\sim}$ for each $p \in \Sigma$. A physical system is said to satisfy the *classical axiom* if each property $a \in \mathscr{L}$ contains a classical question $\gamma \in a$. If the property lattice satisfies this requirement, then $\mathscr{L}$ is order isomorphic with $\mathcal{P}(\Sigma)$, and so the property lattice is a (complete) Boolean algebra. It is sufficient to show that the Cartan morphism $\mu : \mathscr{L} \to \mathcal{P}(\Sigma)$ is surjective when this condition is satisfied.

Proof. Let $p \in \Sigma$, and consider the property $a_p := \bigwedge\{a \in \mathscr{L} \mid p \triangleleft a\}$. If a_p is not an atom of $\mathscr{L}$, there is a property $b \in \mathscr{L}$ with $0 < b < a_p$ and a state $p_1 \in \Sigma$ such that $p_1 \triangleleft b$; let $\beta \in b$ be a classical question. Then $p_1 \triangleleft [\beta]$ and $p \not\triangleleft [\beta]$, hence $p \triangleleft [\beta]^{\sim}$ and so $a_p \leq [\beta^{\sim}]$. Because $b < a_p$ and $p_1 \triangleleft b$ it follows that $p_1 \triangleleft a_p$ and so $p_1 \triangleleft [\beta^{\sim}]$. This is a contradiction and so a_p is an atom of $\mathscr{L}$ and $\mathscr{L}$ is an atomic lattice, with $\mu(a_p) = \{p\}$. Let α be a classical question with $\alpha \in a_p$. If $q \in \Sigma$ and $q \neq p$, then $q \not\triangleleft [\alpha]$ and so $q \triangleleft [\alpha^{\sim}]$, hence $\mu([\alpha^{\sim}]) = \Sigma \setminus \{p\}$.

This concludes the proof, because every subset $A \in \Sigma$ is the intersection of a collection of subsets of this type:

$$A = \bigcap \{\Sigma \setminus \{p\} \mid p \notin A\} . \tag{2.11}$$

□

It turns out that the notion of orthogonality becomes trivial for classical systems, $p \perp q$ if and only if $p \neq q$. For $p \neq q$ there must exist a $\beta \in \mathscr{Q}$ such that $p \triangleleft \beta$ and $q \not\triangleleft \beta$. According to the classical assumption there must exist a classical question $\gamma \in [\beta]$ hence $p \triangleleft \gamma$ so $q \not\triangleleft \gamma$ implies $q \triangleleft \gamma^{\sim}$.

Definition 3. An *observable* is a σ-morphism $\mu : \mathcal{B}(X) \to \mathscr{L}$, where X is a second countable locally compact Hausdorff space, and $\mathcal{B}(X)$ is the σ-algebra of Borel sets in X. Explicitly, μ has the properties

(1) $\mu(X) = I$,

(2) If $E \cap F = \emptyset$, then $\mu(E) \perp \mu(F)$,

[1] See Moore (1999) for more details.

(3) If $n \mapsto E_n$ is a sequence of pairwise disjoint Borel sets,

$$\mu\left(\bigcup_n E_n\right) = \bigvee_n \mu\left(E_n\right). \tag{2.12}$$

That is, an observable consists of a coherent collection of properties. Notice that the range of an observable is a σ-algebra in $\mathscr{L}$.

3 The Galilei and Newton groups

For a classical system the Galilei and Newton symmetry groups are those which are the collection of active or passive transformations respectively, that leave particular physical features invariant. As we will see, they operate at different levels.

The *Galilei group* is a 10-parameter Lie group and is usually defined by its action on the configuration space $\mathbb{R}^4$ as follows:

$$(\mathbf{r}, t) \mapsto (O\mathbf{r} - \mathbf{v}\,t + \mathbf{b}, t + a)\,. \tag{3.1}$$

Here, O is an orthogonal matrix, $\mathbf{v}$ and $\mathbf{a}$ are two vectors, and $a \in \mathbb{R}$. The group composition is given by

$$(R_1, \mathbf{b_1}, \mathbf{w_1}, a_1)\cdot(R_2, \mathbf{b_2}, \mathbf{w_2}, a_2) = (R_1R_2, R_1\mathbf{b}_2 + \mathbf{b}_1 + \mathbf{w}_1a_2, R_1\mathbf{w}_2 + \mathbf{w}_1, a_1 + a_2) \tag{3.2}$$

an inverse

$$(R, \mathbf{b}, \mathbf{w}, a)^{-1} = (R^*, R^*(\mathbf{w}a - \mathbf{b}), -R^*\mathbf{b}, -a), \tag{3.3}$$

and an identity

$$i = (I, \mathbf{0}, \mathbf{0}, 0), \tag{3.4}$$

where I is a 3x3 identity matrix, $\mathbf{0}$ is the origin of $\mathbb{R}^3$ and 0 is the origin of $\mathbb{R}$. The Galilei group is intimately linked to the dynamics of the given system $\mathcal{G}$, because it can be regarded as the dynamical group which leaves invariant the trajectories of the system. The *Newton group* on the other hand is also a 10-parameter group and is defined by its action on $\mathbb{R}^7 = \left\{(\mathbf{q}, \mathbf{p}, t) \mid \mathbf{q}, \mathbf{p} \in \mathbb{R}^3, t \in \mathbb{R}\right\}$ as follows (Giovannini & Piron, 1979):

$$(\mathbf{r}, \mathbf{p}, t) \mapsto (R\mathbf{r} + \mathbf{b}, R\mathbf{p} + \mathbf{w}, t + a)\,. \tag{3.5}$$

Here, $R \in \mathrm{SO}(3)$ and $\mathbf{b}, \mathbf{w} \in \mathbb{R}^3$ and $a \in \mathbb{R}$. Group elements of the Newton group can also be written as $(R, \mathbf{b}, \mathbf{w}, a)$ and the group composition then becomes

$$(R_1, \mathbf{b_1}, \mathbf{w_1}, a_1) \cdot (R_2, \mathbf{b_2}, \mathbf{w_2}, a_2) = (R_1R_2, R_1\mathbf{b}_2 + \mathbf{b}_1, R_1\mathbf{w}_2 + \mathbf{w}_1, a_2 + a_1) \tag{3.6}$$

an inverse

$$(R, \mathbf{b}, \mathbf{w}, a)^{-1} = (R^*, -R^*\mathbf{b}, -R^*\mathbf{w}, -a), \tag{3.7}$$

and an identity

$$i = (I, \mathbf{0}, \mathbf{0}, 0). \tag{3.8}$$

From an operational perspective, this group has some distinct physical advantages. The Newton group arises when a Galilean particle is characterised operationally by how it interacts with the various measurement devices used to probe the system. Therefore, this group acts primarily at the level of the observables that represent these interactions, independent of the dynamics. That is, the Galilean equivalence is implemented at the operational level of the measurement devices, and not directly at the level of the system itself. In other words, if the Galilei group is linked to the dynamics of the system, then the Newton group is associated with coordinate changes between frames of reference. It is for this reason that the Newton group has also been called the *passive Galilei group* (Piron, 1976).

4 The classical elementary particle

One of the simplest non-trivial examples of a classical physical system is the elementary point particle existing in $\mathbb{R}^3$. Using the notions of property and state as discussed in section 2, we start with a detailed exposition of the kinematical structure of the point particle.
The phase space of the point particle is given by

$$\mathbb{R}^6 \cong \mathbb{R}^3 \times \mathbb{R}^3 = \left\{(\mathbf{q}, \mathbf{p}) \mid \mathbf{q}, \mathbf{p} \in \mathbb{R}^3\right\} \tag{4.1}$$

and the corresponding state space is

$$\mathbb{R}^7 \cong \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R} = \left\{(\mathbf{q}, \mathbf{p}, t) \mid \mathbf{q}, \mathbf{p} \in \mathbb{R}^3, t \in \mathbb{R}\right\}. \tag{4.2}$$

The position observable $\mathfrak{q}$ and momentum observable $\mathfrak{p}$ associate the properties of the measurement device (*position* $\mathbf{q}$, *momentum* $\mathbf{p}$ and *time* t in the Borel sets Δ_1, Δ_2 and Δ_3 respectively) to corresponding properties of the classical elementary point particle via the two σ-morphisms to phase space

$$\mathfrak{q} : \mathcal{B}(\mathbb{R}^3) \to \mathcal{B}(\mathbb{R}^6) : \Delta_1 \mapsto \Delta_1 \times \mathbb{R}^3 \tag{4.3}$$

$$\mathfrak{p} : \mathcal{B}(\mathbb{R}^3) \to \mathcal{B}(\mathbb{R}^6) : \Delta_2 \mapsto \mathbb{R}^3 \times \Delta_2 \tag{4.4}$$

or the three mappings to state space

$$\mathfrak{q} : \mathcal{B}(\mathbb{R}^3) \to \mathcal{B}(\mathbb{R}^7) : \Delta_1 \mapsto \Delta_1 \times \mathbb{R}^4 \tag{4.5}$$

$$\mathfrak{p} : \mathcal{B}(\mathbb{R}^3) \to \mathcal{B}(\mathbb{R}^7) : \Delta_2 \mapsto \mathbb{R}^3 \times \Delta_2 \times \mathbb{R} \tag{4.6}$$

$$\mathfrak{t} : \mathcal{B}(\mathbb{R}) \to \mathcal{B}(\mathbb{R}^7) : \Delta_3 \mapsto \mathbb{R}^6 \times \Delta_3. \tag{4.7}$$

These set mappings preserve the lattice operations union, intersection and complementations and so there are essentially unique measurable functions $\hat{q}, \hat{p} : \mathbb{R}^6 \to \mathbb{R}^3$ such that $\hat{q}^{-1}(\Delta) = \mathfrak{q}(\Delta)$ and $\hat{p}^{-1}(\Delta) = \mathfrak{p}(\Delta)$ for all measurable $\Delta \subseteq \mathbb{R}^3$. Specifically, we have

$$\hat{q}((\mathbf{q}, \mathbf{p})) = \mathbf{q} \tag{4.8}$$

$$\hat{p}((\mathbf{q}, \mathbf{p})) = \mathbf{p} \tag{4.9}$$

In addition, any other observable into $\mathcal{B}(\mathbb{R}^6)$ is then a function of these two given observables. The Newton group acts on the two copies of $\mathbb{R}^3$ associated with the measurement devices for the two types of observables in slightly different ways; it is sufficient to consider the various rotations and spatial translations $\vec{a}$ and boosts $\vec{v}$: We have $\vec{a}(\Delta) = \Delta + \mathbf{a}$, $\vec{v}(\Delta) = \Delta$ on the copy of $\mathcal{B}(\mathbb{R}^3)$ associated with the position observable, and $\vec{a}(\Delta) = \Delta$ and $\vec{v}(\Delta) = \Delta + \mathbf{v}$ on the copy of $\mathcal{B}(\mathbb{R}^3)$ associated with the momentum observable.

The corresponding systems of imprimitivity are then given by the diagrams

$$\begin{array}{ccc} \mathcal{B}(\mathbb{R}^3) & \xrightarrow{\mu} & \mathcal{B}(\mathbb{R}^6) \\ {\scriptstyle R}\downarrow & & \downarrow{\scriptstyle S(R)} \\ \mathcal{B}(\mathbb{R}^3) & \xrightarrow[\mu]{} & \mathcal{B}(\mathbb{R}^6) \end{array} \tag{4.10}$$

$$\begin{array}{ccc} \mathcal{B}(\mathbb{R}^3) & \xrightarrow{\mu} & \mathcal{B}(\mathbb{R}^6) \\ {\scriptstyle \vec{a}}\downarrow & & \downarrow{\scriptstyle S(\vec{a})} \\ \mathcal{B}(\mathbb{R}^3) & \xrightarrow[\mu]{} & \mathcal{B}(\mathbb{R}^6) \end{array} \tag{4.11}$$

$$\begin{array}{ccc} \mathcal{B}(\mathbb{R}^3) & \xrightarrow{\quad\mu\quad} & \mathcal{B}(\mathbb{R}^6) \\ {\scriptstyle \vec{v}}\downarrow & & \downarrow{\scriptstyle S(\vec{v})} \\ \mathcal{B}(\mathbb{R}^3) & \xrightarrow[\quad\mu\quad]{} & \mathcal{B}(\mathbb{R}^6) \end{array} \tag{4.12}$$

where μ is either $\mathfrak{q}$ or $\mathfrak{p}$; there is a similar diagram for $\mathfrak{t}$ on state space, where translations only act on the domain space. For instance, the first two diagrams for the localisation observable $\mathfrak{q}$ state that the system is localised in the subset $gE \subseteq \mathbb{R}^3$ iff the property $S(g)(\mathfrak{q}E)$ is actual. More generally, if s denotes an arbitrary state of the system, we have

$$s(\mathfrak{q}(gE)) = s(S(g)(\mathfrak{q}E)\,. \tag{4.13}$$

The corresponding actions on phase and state spaces are then given by

$$S(R)(\mathbf{q},\mathbf{p}) = (R\mathbf{q}, R\mathbf{p}) \tag{4.14}$$

$$S(\vec{a})(\mathbf{q},\mathbf{p}) = (\mathbf{q}+\mathbf{a}, \mathbf{p}) \tag{4.15}$$

$$S(\vec{v})(\mathbf{q},\mathbf{p}) = (\mathbf{q}, \mathbf{p}+\mathbf{v}) \tag{4.16}$$

and

$$S(R)(\mathbf{q},\mathbf{p},t) = (R\mathbf{q}, R\mathbf{p}, t) \tag{4.17}$$

$$S(\vec{a})(\mathbf{q},\mathbf{p},t) = (\mathbf{q}+\mathbf{a}, \mathbf{p}, t) \tag{4.18}$$

$$S(\vec{v})(\mathbf{q},\mathbf{p},t) = (\mathbf{q}, \mathbf{p}+\mathbf{v}, t) \tag{4.19}$$

$$S(\tau)(\mathbf{q},\mathbf{p},t) = (\mathbf{q}, \mathbf{p}, t+\tau) \tag{4.20}$$

respectively. One easily verifies that these actions on phase space $S : G \times \mathcal{B}(\mathbb{R}^6) \to \mathcal{B}(\mathbb{R}^6)$, with a similar action on state space, satisfy the commutativity requirements above. For instance, we have

$$(S(R) \circ \mathfrak{q})(\Delta) = S(R)(\Delta \times \mathbb{R}^3) = R(\Delta) \times \mathbb{R}^3 = (\mathfrak{q} \circ R)(\Delta)\,. \tag{4.21}$$

In other words, the usual textbook mathematical model for the classical point particle satisfies all the natural symmetry requirements that arise from our operational perspective.

5 Group actions and the stabiliser subgroup

In order to introduce systems of imprimitivity in the next section, some background theory with group actions is required, stabiliser subgroups in particular.

Let G be a group with a left action on a set X. The *orbit* Gx of a point $x \in X$ is the collection of every point in X which can be mapped to by the action

$$Gx = \{gx \mid g \in G\}. \tag{5.1}$$

An action is considered *transitive* if the whole of X is the orbit of some point $x \in X$

$$\exists x \in X \text{ such that } X = Gx. \tag{5.2}$$

The *stabiliser subgroup* of the point x is the subgroup $G_x \subseteq G$ which contains all the transformations on x which leave x invariant.

$$G_x = \{g \in G \mid gx = x\} \tag{5.3}$$

One verifies easily that this is indeed a subgroup of G, for each point x.

It will be useful to mention that there is a G-equivarient isomorphism between any transitive action of a group G on a set X (i.e. $gx = x$) is equivariant with the natural action of G on the coset space G/G_x. First, we show that all stabiliser subgroups for points in X are isomorphic.

Proposition 2. *Assume that the group G acts transitively on the set X, and let $x_0, x_1 \in X$. If $\tilde{g} \in G$ with $\tilde{g}x_0 = x_1$, then $G_{x_1} = \tilde{g}G_{x_0}\tilde{g}^{-1}$.*

Proof. For $g \in G_{x_1}$ and $x_0, x_1 \in X = Gx$,

$$(\tilde{g}^{-1}g\tilde{g})x_0 = \tilde{g}^{-1}gx_1 \tag{5.4}$$

$$= \tilde{g}^{-1}x_1 \tag{5.5}$$

$$= x_0 \tag{5.6}$$

This shows that $\tilde{g}^{-1}g\tilde{g} \in G_{x_0}$ and thus $\tilde{g}^{-1}G_{x_1}\tilde{g} \subseteq G_{x_0}$. The same approach can be used to show $\tilde{g}G_{x_0}\tilde{g}^{-1} \subseteq G_{x_1}$ and so $G_{x_0} \subseteq \tilde{g}^{-1}G_{x_1}\tilde{g}$, from which equality follows. □

The following result establishes the fact that it is sufficient to consider transitive actions of a given group G on coset spaces for that group.

Theorem 1. Each transitive G-action is isomorphic to the action of G by left translation on a coset space G/H, for some subgroup H of G.

Proof. Choose a point $x_0 \in X$, and construct a mapping

$$f : G \to X : g \mapsto gx_0 \,.$$

G being transitive, this map is surjective. Now $g_1 x_0 = g_2 x_0$ iff $g_1^{-1} g_2 \in S_{x_0}$, hence there exists a unique and well-defined bijection $\phi : G/S_{x_0} \to X$ given by $\phi(gS_{x_0}) = f(g)$. The mapping ϕ is G-equivariant, because

$$\phi(g_1 g_2 \, S_{x_0}) = f(g_1 g_2) = g_1 g_2 x_0 = g_1 f(g_2) = g_1 \phi(g_2 \, S_{x_0})$$

which is exactly what we need. Because ϕ is a G-equivariant bijection, it is a G-isomorphism. □

6 Systems of imprimitivity

From a group theoretical perspective, the structure of a group can be studied by constructing representations of the group on various mathematical structures. In particular, classical group representation theory constructs copies of a given group in the general linear group acting on some vector space $\mathrm{GL}(V)$; more precisely, a representation of G is a group morphism $G \to \mathrm{GL}(V)$. A representation is said to be *irreducible* if the carrier vector space V has no proper subspaces invariant under the action of G.

In a similar way, we can investigate the structure of a group by constructing permutation representations of the groups on sets, perhaps with additional structure. The general concept of a group action $G \times X \to X$ with X some set is doing exactly this. Equivalently, a group action on the set X is a group morphism $G \to \mathrm{Sym}(X)$, where the latter set denotes the collection of all permutations of X . In this spirit, a group action can be regarded as a *permutation representation* of the group G.

To make the connection with the "standard model" in the previous section, we will use the fact that any transitive action of a group G on a set X is equivariant with the natural action of G on the coset space G/G_x. This result can be regarded as a special case of the general theory of systems of imprimitivity applied to permutation representations. As we will see, the physical concept of a *covariant* observable corresponds with the mathematical concept of a system of imprimitivity, at least when the space in which the observable takes its values is a transitive G-space.

It is important to note that the following explanation is only a toy model and makes use of finite sets rather than realistic models that would have additional topological and measure theoretic structure, this method however gives us a clearer understanding as to isolate the structural properties of the concepts of interest.

Let S be an H-space, where H is a subgroup of G. It turns out that we can always construct an essentially unique larger G-set in a purely syntactical way, which preserves the action of H on S. The construction proceeds as follows:

Start with the cartesian product $G \times S$. Each ordered pair (g, s) can be interpreted formally as the group element $g \in G$ acting on $s \in S$, and in order for this formal action of G to be compatible with the given action H we have to require that, for all $g \in G$, $h \in H$ and $s \in S$

$$(gh, s) = (g, hs)\,. \tag{6.1}$$

This requirement induces an equivalence relation on $G \times S$. We say that $(g_1, s_1) \sim (g_2, s_2)$ if there is an $h \in H$ such that $g_1 = g_2 h$ and $s_2 = hs_1$; that is, $(gh, s) \sim (g, hs)$. The usual procedure yields the quotient set, which we will denote by $G \times_H S$. The natural action of G on $G \times S$ (given by $g * (g_1, s_1) = (gg_1, s_1)$) is compatible with the equivalence relation:

$$g * (g_1 h, s) = (gg_1 h, s) \sim (gg_1, hs) = g * (g_1, hs) \tag{6.2}$$

and so this action descends into the quotient which then also becomes a G-set with action

$$g \star [(g_1, s)] = [(gg_1, s)]. \tag{6.3}$$

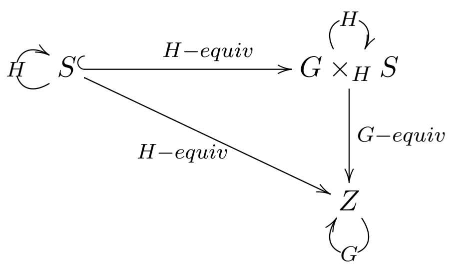
There is a natural H-equivariant injection $j : S \to G \times_H S$:

$$j(s) = [(e, s)] \tag{6.4}$$

where e is the group identity. We have indeed

$$j(hs) = [(e, hs)] = [(h, s)] = h \star [(e, s)] = h \star j(s) \tag{6.5}$$

and so we can identify the H-set S with the subset $\{[(e, s)] \mid s \in S\}$ which is left invariant by the subgroup H of G. From this perspective, we have extended the known action of the subgroup H on S into an action of the larger group G onto a larger set. In fact, the extension above is essentially unique; this is the content of the second part in the following statement. The following diagram summarises these concepts.



(6.6)

Theorem 2. With S, H, G and $j : S \to G \times_H S$ as above we have

(1) $Gj(S) = G \times_H S$.

(2) If Z is a G-set and $k : S \to Z$ is H-equivariant, there is a unique G-equivariant map $\tilde{k} : G \times_H S \to Z$ such that $\tilde{k} \circ j = k$.

The G-set $G \times_H S$ is said to be a *free* G-extension of the H-set S. It is essentially unique in the sense that any other G-extension satisfying these two properties is G-equivalent with $G \times_H S$. Notice that this extension procedure is purely syntactical, in the sense that all the interesting mathematical content resides in the action of H on G.

How does all this relate to our notion of covariant observables? First, we need the following

Definition 4. Let X be a G-set. A partition $\mathcal{E}$ of X is called a *system of imprimitivity* if G acts transitively on $\mathcal{E}$.[2]

Here, G acts on subsets of X in the usual way: If $E \subseteq X$ then $gE := \{ge \mid e \in E\}$. In other words, for each $E \in \mathcal{E}$, $gE \in \mathcal{E}$, this action is transitive and, for each $E \in \mathcal{E}$ we have $\bigcup \{gE \mid g \in G\} = X$. Each of the sets in $\mathcal{E}$ is called a *set of imprimitivity*. We then have the following *Imprimitivity Theorem.*

Theorem 3. Let X be a G-set, and let $S \subseteq X$ with stabiliser $H := G_s$

$$G_s = \{g \in G \mid gS \subseteq S\}. \tag{6.7}$$

The following are equivalent:

(1) The subset S generates a system of imprimitivity $\{gS \mid g \in G\}$.

(2) There exists a G-equivalence $f : G \times_H S \to X$.

It is the second part of the theorem which is important for our purposes. In a nutshell, let $F : \mathcal{B}(\mathcal{R}^3) \to \mathcal{P}(\Sigma)$ be a covariant observable relative to the group G. One can show that there exists a G-equivariant function $f : \Sigma \to \mathbb{R}^3$ such that $F = f^{-1}$. If the action of G on $\mathcal{B}(\mathcal{R}^3)$ is transitive then f is surjective. It follows that $S := f^{-1}(m_0)$ is a set of imprimitivity invariant under the action of $H := G_s$ and so there exists a G-equivalence between Σ and $G \times_H S$. It is in this sense that the structure of the state space Σ is determined in an essentially unique way by the properties of the symmetry group.

[2]Recall that a partition of a set consists of a collection of pairwise disjoint subsets that cover X.

References

Birkhoff & von Neumann. 1936. The logic of quantum mechanics. Annals of Mathematics **37**: 823–843.

Foulis & Wilce. 2000. Free extensions of group actions, induced representations, and the foundations of physics. In Coecke & Wilce. 2000. *Current Research in Operational Quantum Logic – Algebras, Categories, Languages.* Kluwer Academic Publishers.

Giovannini & Piron. 1979. On the group-theoretical foundations of classical and quantum physics: kinematics and state spaces. Helv. Phys. Acta **52**: 518–540.

Jauch. 1968. *Foundations of Quantum Mechanics.* Addison-Wesley.

Kibble & Berkshire. 2004. *Classical Mechanics.* Imperial College Press.

Lévy-Leblond, 1971. Galilei Group and Galilean Invariance. In Loebl. 1971. *Group Theory and its Applications.* Volume 3. Academic Press.

Loomis & Sternberg. 1968. *Advanced Calculus.* Addison-Wesley.

Moore. 1999. On State Spaces and Property Lattices. Stud. Hist. Phil. Mod. Phys. **30**: 61–83.

Piron. 1976. *Foundations of Quantum Physics.* W. A. Benjamin.

Piron. 1998. *Mécanique quantique. Bases et applications.* Presses polytechniques et universitaires romandes.

Szekeres. 2004. *A Course in Modern Mathematical Physics.* Cambridge University Press.