Representation Theory and Quantum Mechanics

By:
Justin Kulp

Faculty Advisor:
Dr. Andrew J. Dean

A project submitted to the Department of Mathematical Sciences in conformity with the requirements for Math-4301 (Honours Seminar)

Lakehead University
Thunder Bay, Ontario

Abstract

In this paper we establish the mathematics used in quantum mechanics with a brief overview of functional analysis and an introduction to the representation theory of Lie groups. We then develop quantum mechanics in a mathematically rigorous way and briefly investigate simple applications of representation theory in modelling physical phenomena.

Contents

0 The History of Quantum Mechanics

> "*The more important fundamental laws and facts of physical science have all been discovered, and these are now so firmly established that the possibility of their ever being supplanted in consequence of new discoveries is exceedingly remote... Our future discoveries must be looked for in the sixth place of decimals.*"
>
> Albert A. Michelson, *Light Waves and Their Uses*, 1902

The subject known as classical physics can be said to have started with the publication of Sir Isaac Newton's 1687 volumes *Philosophiæ Naturalis Principia Mathematica*. Until the start of the 20th century, the mechanics of Newton, Lagrange, and Hamilton; the electrodynamics of Maxwell; the application of statistics to large systems; and Einstein's relativity had sufficed to describe almost all known phenomenon. As Albert Michelson commented, physics had been believed to have been reduced to a few small open issues, and precision and accuracy problems.

However, experimental results on atomic phenomenon led to the realization that classical physics was not adequate very soon after. Problems with divergent integrals in thermal physics and electromagnetism led to predictions suggesting that certain objects could radiate infinite energy. As well, measurements on atomic phenomenon (on the scale of $\approx 10^{-10}m$) suggested that the mechanics of the very small was not exactly deterministic in the usual sense.

Solutions to many of the open problems at the time were simultaneously resolved with the development of quantum mechanics. Originally, it began with the work of Bohr, Planck, and others, who developed ad hoc explanations and rules to describe individual phenomena in a highly informal way between 1900 and 1925. Soon after, coherent descriptions of quantum mechanics were developed by Schrödinger and Heisenberg which are the basis of that used today. The mathematically rigorous descriptions of quantum mechanics that we will discuss evolved from the work of von Neumann, Dirac, Weyl, and Wigner [9].

Today, relativistic and non-relativistic quantum mechanics and its foundations are still active areas of research for mathematicians, physicists, and computer scientists, in particle theory, quantum computing, quantum optics, and quantum materials.

1 Mathematical Preliminaries

Here we will extend concepts from linear algebra, group theory, and analysis necessary for our later discussions on both mathematics and physics. If a ground field is not specified it is $\mathbb{C}$.

1.1 Multilinear Algebra

We begin by assigning some terminology to familiar concepts from linear algebra and abstract algebra. The terminology and definitions are from [2] and [3].

Let V and W be vector spaces over a field $\mathbb{F}$. If $\phi : V \to W$ preserves both vector addition and scalar multiplication then ϕ is a homomorphism. The vector space of all homomorphisms from V to W is denoted $\mathrm{Hom}_{\mathbb{F}}(V, W)$. If ϕ is a bijection then it is called an isomorphism. If ϕ is a homomorphism and $W = V$ then ϕ is an endomorphism on V. The ring of all endomorphisms is $\mathrm{End}_{\mathbb{F}}(V)$. If ϕ is both an isomorphism and $W = V$, then ϕ is an automorphism of V. The group of all automorphisms of V is $\mathrm{GL}(V)$.

Definition 1.1. *Let V be a vector space over $\mathbb{F}$. $V^* = \mathrm{Hom}_{\mathbb{F}}(V, \mathbb{F})$ is the dual space of V, and it's elements are called linear functionals.*

If V is finite dimensional, $\dim V = \dim V^*$. If V is a finite dimensional space of column vectors we can view V^* as the space of row vectors.

Definition 1.2. *Let $V_1, \ldots, V_n$ and W be vectors spaces. A function $F : V_1 \times \cdots \times V_n \to W$ is multilinear if $F(v_1, \ldots, v_n)$ is linear in each v_i when all other elements are fixed. F is called bilinear when $n = 2$. F is alternating if permuting any two elements adds a factor of -1. F is symmetric if permuting any two elements leaves F unchanged.*

Definition 1.3. *Let V and W be two vector spaces over $\mathbb{F}$. The tensor product $V \otimes_{\mathbb{F}} W$ is a new vector space with a universal bilinear map $\psi : V \times W \to V \otimes_{\mathbb{F}} W$ mapping $v \times w \to v \otimes w$. ψ is universal in the sense that if $\beta : V \times W \to U$ is any bilinear map from $V \times W$ to a vector space U, then there exists a unique linear map $\tilde{\beta} : V \otimes_{\mathbb{F}} W \to U$. We often omit the subscript on $\otimes$.*

Proposition 1.4. *Let V, W, and U be vector spaces. Say $\{v_i\}$ and $\{w_i\}$ are bases for V and W. Then the following hold:*

1. $V \otimes W \cong W \otimes V$
2. $(V \oplus W) \otimes U \cong (V \otimes U) \oplus (W \otimes U)$

3. $(V \otimes W) \otimes U \cong V \otimes (W \otimes U) \cong V \otimes W \otimes U$
4. *A basis for* $V \otimes W$ *is* $\{v_i \otimes w_i\}$
5. $\dim(V \otimes W) = (\dim V)(\dim W)$

Definition 1.5. *Let* V *be a vector space. The* n^{th} *exterior power is a vector space* $\bigwedge^n V$ *with a universal alternating multilinear map* $\psi : V^n \to \bigwedge^n V$ *mapping* $v_1 \times \cdots \times v_n \to v_1 \wedge \cdots \wedge v_n$. *The* n^{th} *symmetric power is a vector space* $\mathrm{Sym}^n V$ *with a universal symmetric multilinear map* $\psi : V^n \to \mathrm{Sym}^n V$ *mapping* $v_1 \times \cdots \times v_n \to v_1 \cdots v_n$.

We may speak of a vector in a tensor power $V^{\otimes n}$ as being totally anti-symmetric or totally symmetric, that is, it behaves like an element of $\bigwedge^n V$ or $\mathrm{Sym}^n V$. We may speak safely about $\bigwedge^n V$ and $\mathrm{Sym}^n V$ inside $V^{\otimes n}$ because there are natural imbeddings $\mathfrak{i}_\wedge : \bigwedge^n V \hookrightarrow V^{\otimes n}$ and $\mathfrak{i}_S : \mathrm{Sym}^n V \hookrightarrow V^{\otimes n}$ given by

$$\mathfrak{i}_\wedge(v_1 \wedge \cdots \wedge v_n) = \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}$$

and

$$\mathfrak{i}_S(v_1 \cdots v_n) = \sum_{\sigma \in \mathfrak{S}_n} v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(n)}$$

Definition 1.6. *Let* V *and* W *be vector spaces. A vector in* $V \otimes W$ *is decomposable if it can be written* $v \otimes w$ *for some* $v \in V$ *and* $w \in W$. *Otherwise it is called indecomposable.*

When we are working with a specific basis we have to be sure that no change of basis can transform a vector that looks indecomposable into one of the form $v \otimes w$.

1.2 Measure Theory and Functional Analysis

1.2.1 Measures

Here develop some basic definitions in measure theory. If X is some set, let $\wp(X)$ denote its power set, and $\setminus$ denote complements. For more information see [12].

Definition 1.7. *Let* X *be a set and* $\tau \in \wp(X)$. *A topological space is a pair* (X, τ) *such that*

1. $\emptyset \in \tau$ *and* $X \in \tau$
2. *If* $\{E_\alpha\}$ *is a finite collection of members of* τ, *then* $\bigcap E_\alpha \in \tau$
3. *If* $\{E_\alpha\}$ *is a ANY collection of members of* τ, *then* $\bigcup E_\alpha \in \tau$

Definition 1.8. *Let* X *be a set and* $\Sigma \in \wp(X)$. Σ *is a* σ*-algebra over* X *if*

1. $X \in \Sigma$
2. *If* $E \in \Sigma$, *then* $X \setminus E \in \Sigma$
3. *If* $\{E_\alpha\}$ *is a countable collection of members of* Σ, *then* $\bigcup E_\alpha \in \Sigma$

τ is called a topology in X, and the members of τ are the open sets in X. There exists a smallest σ-algebra, $\mathscr{B}$, containing every open set in X, these are the Borel sets of X.

Definition 1.9. *Let* X *be a set,* Σ *a* σ*-algebra over* X. $\mu : \Sigma \to [0, \infty]$ *is a measure and* (X, Σ, μ) *is a measure space if they satisfy*

1. $\mu(\emptyset) = 0$
2. $\mu(E) \geq 0$ *for all* $E \in \Sigma$
3. *If* $\{E_\alpha\}$ *is a pairwise disjoint countable collection of members of* Σ, *then*

$$\mu\left(\bigcup E_\alpha\right) = \sum \mu(E_\alpha)$$

1.2.2 Self-Adjoint Operators on Hilbert Space

We conclude with a discussion of the basics of Hilbert spaces and self-adjoint operators. While this is an extremely large subject we will only present the absolute minimum for our discussion. Proofs of the long and non-immediate theorems can be found in [7] and [10].

If V and W are normed spaces, a linear map $T : V \to W$ is bounded if the norm is finite, where the norm is defined as

$$\|T\| \coloneqq \sup\{\|T\psi\| : v \in V, \|v\| = 1\}$$

By construction $\|T\|$ is the least number such that $\|Tv\| \leq \|T\|\|v\|$ for all $v \in V$.

Proposition 1.10. *Let* V *and* W *be normed vector spaces. Let* $T : V \to W$ *be a linear transformation. The following are equivalent:*

1. T *is bounded.*
2. T *is continuous.*
3. T *is continuous at one point of* V.

Proof. By $\|T(v_1 - v_2)\| \leq \|T\|\|v_1 - v_2\|$, (1) $\implies$ (2) since Lipschitz functions are continuous. (2) $\implies$ (3) by definition. Suppose that T is continuous at some v_0, so that if $\|v\| = \|(v + v_0) - v_0\| < \delta$ then $\|Tv\| = \|T(v + v_0) - T(v_0)\| < \epsilon$, then $\|T\| \leq \epsilon/\delta$, so (3) $\implies$ (1) $\square$

Definition 1.11. *A Hilbert space, $\mathcal{H}$, is a real or complex vector space with inner product $\langle\cdot|\cdot\rangle$, which is complete as a metric space with respect to the norm $\|\psi\| = \sqrt{\langle\psi|\psi\rangle}$. $\mathcal{H}$ is called separable if it has a countable orthonormal basis.*

From here out we assume $\mathcal{H}$ is complex, we use the convention that $\langle\phi|\psi\rangle$ is anti-linear in the first slot, so it satisfies

1. $\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$
2. $\langle\phi|\xi_1\psi_1 + \xi_2\psi_2\rangle = \xi_1\langle\phi|\psi_1\rangle + \xi_2\langle\phi|\psi_2\rangle$
3. $\langle\eta_1\phi_1 + \eta_2\phi_2|\psi\rangle = \eta_1^*\langle\phi_1|\psi\rangle + \eta_2^*\langle\phi_2|\psi\rangle$
4. $\langle\psi|\psi\rangle \geq 0$, with equality $\iff \psi = 0$

Theorem 1.12 (The Riesz Representation Theorem)**.**
If ϕ^* is a bounded linear functional of $\mathcal{H}$, then there is a unique $\phi \in \mathcal{H}$ such that

$$\phi^*(\psi) = \langle\phi|\psi\rangle \quad \text{for all } \psi \in \mathcal{H}$$

Proposition 1.13. *Let $\mathcal{B}(\mathcal{H})$ denote the set of bounded linear maps on $\mathcal{H}$. For any $T \in \mathcal{B}(\mathcal{H})$ there is a unique linear operator $T^\dagger : \mathcal{H} \to \mathcal{H}$ such that for any $\phi, \psi \in \mathcal{H}$*

$$\langle\phi|T\psi\rangle = \langle T^\dagger\phi|\psi\rangle$$

If $T = T^\dagger$ then T is called self-adjoint or Hermitian.

If $\mathcal{H}$ is finite dimensional, the linear operators are matrices, and the Hermitian operators are the matrices that are their complex conjugate transpose, $A = A^\dagger = (A^*)^T$.

Definition 1.14. *A linear operator U on $\mathcal{H}$ is unitary if it is surjective and preserves the inner product of $\mathcal{H}$, $\langle U\phi|U\psi\rangle = \langle\phi|\psi\rangle$.*

Proposition 1.15. *A bounded linear operator U is unitary if and only if $U^\dagger = U^{-1}$.*

Definition 1.16. *A one-parameter unitary group on $\mathcal{H}$ is a collection of unitary operators $U(t)$, $t \in \mathbb{R}$, with the property that $U(0) = \mathbb{I}$ and $U(s+t) = U(s)U(t)$ for all $s, t \in \mathbb{R}$. It is strongly continuous if for all $\psi \in \mathcal{H}$ and $t \in \mathbb{R}$ we have*

$$\lim_{s\to t}\|U(t)\psi - U(s)\psi\| = 0$$

Theorem 1.17 (Stone's Theorem)**.**
Suppose $U(\cdot)$ is a strongly continuous one-parameter unitary group on $\mathcal{H}$, then for some self-adjoint operator A on $\mathcal{H}$ we have $U(t) = e^{itA}$ for all $t \in \mathbb{R}$.

Definition 1.18. *If $A \in \mathcal{B}(\mathcal{H})$ the resolvent set, $\rho(A)$, is all $\lambda \in \mathbb{C}$ such that $(A - \lambda\,\mathbb{I})^{-1}$ is bounded. The spectrum of A, $\sigma(A)$, is defined to be $\mathbb{C} \setminus \rho(A)$.*

For any closed subspace $E \subseteq \mathcal{H}$, there is a unique bounded operator P_E such that $P_E = \mathbb{I}$ on E and $P_E = 0$ on $E^{\perp}$. This operator is the *orthogonal projection* onto E.

Definition 1.19. *Let X be a set and Σ a σ-algebra over X. A map $\mu : \Sigma \to \mathcal{B}(\mathcal{H})$ is called a projection-valued measure if*

1. *$\mu(\emptyset) = 0$ and $\mu(X) = \mathbb{I}$*
2. *For each $E \in \Sigma$, $\mu(E)$ is an orthogonal projection*
3. *If $\{E_\alpha\}$ is a pairwise disjoint countable collection of members of Σ, then*

$$\mu\left(\bigcup E_\alpha\right)\psi = \sum \mu(E_\alpha)\psi$$

Theorem 1.20 (The Spectral Theorem)**.**
If $A \in \mathcal{B}(\mathcal{H})$ is self-adjoint, then there exists a unique projection-valued measure, μ^A, on the Borel σ-algebra in $\sigma(A)$, with values in projections on $\mathcal{H}$, such that

$$A = \int_{\sigma(A)} \lambda \; d\mu^A(\lambda)$$

Proposition 1.21. *Let $A \in \mathcal{B}(\mathcal{H})$ be a self-adjoint operator on $\mathcal{H}$. Then the eigenvalues of A are real. Eigenvectors of A corresponding to different eigenvalues are orthogonal.*

Proof. Suppose $A\psi = \lambda\psi$, by hypothesis $\langle\psi|A\psi\rangle = \langle A\psi|\psi\rangle$, so $\lambda\langle\psi|\psi\rangle = \lambda^*\langle\psi|\psi\rangle$. Eigenvectors must be non-zero be definition hence $\lambda = \lambda^*$ and λ is real. Now suppose $A\psi' = \lambda'\psi'$, then $\langle\psi|A\psi'\rangle = \langle A\psi|\psi'\rangle$ implies $\lambda^*\langle\psi|\psi'\rangle = \lambda'\langle\psi|\psi'\rangle$. Since $\lambda^* = \lambda \neq \lambda'$, then we must have $\langle\psi|\psi'\rangle = 0$. □

Definition 1.22. *Let $0 \neq \Psi \in \mathcal{H}$, the ray containing Ψ is the set $\mathbf{\Psi} = \{e^{i\alpha}\psi \mid \alpha \in \mathbb{R}\}$. If $\|\psi\| = 1$ then $\mathbf{\Psi}$ is a unit ray.*

2 Lie Groups and Lie Algebras

In this section we look at the definitions of Lie groups and Lie algebras and their representations, following [3], [5], and [8]. We also investigate the irreducible representations of the most important Lie algebra, $\mathfrak{sl}_2\mathbb{C}$.

2.1 Lie Groups and Lie Algebras

2.1.1 Manifolds

Suppose (X, τ) is a topological space. Then X is T_2 if for any $x, y \in X$ there are neighbourhoods U_x and U_y, of x and y respectively, such that $U_x \cap U_y = \emptyset$. A *base* of the space is a subset $B \subseteq \tau$ such that every open set in τ is a union of elements in B. A *homeomorphism* is a continuous function between topological spaces with a continuous inverse.

Definition 2.1. *Let (X, τ) be a T_2 space with a countable basis for τ. An n-chart for X is a pair, (U, Φ), with U an open subset in X, and $\Phi : U \to \mathbb{R}^n$ a homeomorphism of U onto the open set $\Phi(U) \subseteq \mathbb{R}^n$.*

Definition 2.2. *Let (X, τ) be a T_2 space with a countable basis, a C^∞ n-atlas for X is a collection $\mathcal{A} = \{(U_\alpha, \Phi_\alpha)\}$ of n-charts for X such that*

1. The open sets $\{U_\alpha\}$ cover X

2. For any indices α, β, the maps $\Phi_\beta \circ \Phi_\alpha^{-1} : \Phi_\alpha(U_\alpha \cap U_\beta) \to \Phi_\beta(U_\alpha \cap U_\beta)$ are C^∞.

An n-chart (U, Φ) is compatible with $\mathcal{A}$ if whenever $U \cap U_\alpha \neq \emptyset$ then the maps $\Phi \circ \Phi_\alpha^{-1}$ and $\Phi_\alpha \circ \Phi^{-1}$ between open sets $\Phi_\alpha(U \cap U_\alpha)$ and $\Phi(U \cap U_\alpha)$ are C^∞.

An *n-dimensional C^∞ structure* on X is an n-atlas, $\mathcal{A}$, for X, such that every n-chart of X compatible with $\mathcal{A}$ is contained in $\mathcal{A}$. If $\mathcal{A}$ is a C^∞ n-atlas for X, then $\mathcal{A}$ is contained in a unique n-dimensional C^∞ structure for X.

Definition 2.3. *A pair $M = (X, \mathcal{A})$ of a T_2 topological space with countable basis, X, and an n-dimensional C^∞ structure, $\mathcal{A}$, on X is called a real C^∞ manifold.*

When we refer to a point in M, we mean a point in X. If we define the n-charts so that $\Phi : U \to \mathbb{C}^n$, M is a *complex C^∞ manifold*. Let $M = (X, \mathcal{A})$ and $N = (Y, \mathcal{B})$ be C^∞ manifolds. A *C^∞ map* $f : M \to N$ is a continuous map $f : X \to Y$ such that whenever $(U, \Phi) \in \mathcal{A}$ and $(B, \Psi) \in \mathcal{B}$ satisfy $U \subseteq f^{-1}(V)$, then $\Psi \circ f \circ \Phi^{-1}$ is C^∞. The C^∞ maps on a manifold M is $C^\infty(M)$.

Definition 2.4. *Let M an n-dimensional C^∞ manifold. A point derivation, L, of $C^\infty(M)$ at $x \in M$ is a linear functional on $C^\infty(M)$ such that*

$$L(fg) = f(x)L(g) + g(x)L(f) \quad \forall f, g \in C^\infty(M)$$

The vector space of point derivations at x is the tangent space to M at x, $T(M)_x$

Definition 2.5. *Let M and N be C^∞ manifolds. If $f : M \to N$ is C^∞, the differential of f at x is the map $df_x : T(M)_x \to T(N)_{f(x)}$ given by*

$$df_x(L)g = L(g \circ f) \quad \text{for all } L \in T(M)_x \text{ and } g \in C^\infty(M)$$

A *vector field*, V, is a map assigning each $x \in M$ a vector in its tangent plane $T(M)_x$. We denote this vector at x by $V(x)$. Suppose V is a vector field on a manifold M, and $x_0 \in M$, an *integral curve* for V with initial condition x_0 is a curve $\gamma : (a, b) \subseteq \mathbb{R} \to M$ such that $\gamma'(t) = V(\gamma(t))$ for every $t \in (a, b)$, $0 \in (a, b)$, and $\gamma(0) = x_0$. There is a unique maximal integral curve for V with initial condition x_0.

Definition 2.6. *Let M and N be C^∞ manifolds and suppose $N \subseteq M$ as a set, but not necessarily as a topological subspace. Let $l : N \to M$ by $x \to x$. N is a submanifold of M if l is C^∞ and dl_x is injective.*

2.1.2 Lie Groups

Definition 2.7. *A Lie group, G, is a group that is also a C^∞ manifold, with both the group operations $* : G \times G \to G$ and $^{-1} : G \to G$ being C^∞.*

By identification of the n^2 entries in an element of $M_n(\mathbb{F})$ with a vector in $\mathbb{F}^{n^2}$, we have that $M_n(\mathbb{F})$ is a T_2 space with a countable basis.

Definition 2.8. *If G and H are Lie groups, and H is a subgroup and submanifold of G, then H is a Lie subgroup of G.*

Definition 2.9. *A Lie group homomorphism, $\rho : G \to H$, is a group homomorphism from G to H which is also a C^∞ map.*

Proposition 2.10. *$GL_n(\mathbb{R})$ and $GL_n(\mathbb{C})$ are open subsets of $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$ respectively. Also, $GL_n(\mathbb{R})$ and $GL_n(\mathbb{C})$ (and their Lie subgroups) are Lie groups.*

Proof. We will prove it just in the case of $\mathbb{C}$, the proof of $\mathbb{R}$ is identical. Consider the determinant map, $\det : \mathrm{GL}_n \mathbb{C} \to \mathbb{C}^*$. Since it is a polynomial function in the

matrix elements, det is continuous. Since $\mathbb{C}^*$ is an open set in $\mathbb{C}$, $GL_n(\mathbb{C})$ is an open set in $M_n(\mathbb{C})$. Note in the real case that $M_n(\mathbb{R})$ is disconnected.

Let $X, Y \in GL_n(\mathbb{C})$. Then XY is a polynomial in the (variable) entries of X and Y. So group multiplication is C^∞. The inverse of some $X \in GL_n(\mathbb{C})$ is also a polynomial in the entries of X, so the inverse operation is C^∞. □

$SO(n, \mathbb{R})$ is the group of rotations about the origin in n-dimensional Euclidean space. If we choose an orthonormal basis this group can be described by those n-dimensional matrices such that $A^T A = \mathbb{I}$ and $\det(A) = 1$. Define $g : GL_n(\mathbb{R}) \to GL_n(\mathbb{R})$ by $g(A) = det(A)A^T A$, as in the proposition, g is continuous and $g^{-1}(\{\mathbb{I}\}) = SO(n, \mathbb{R})$, so $SO(n, \mathbb{R})$ is a connected Lie group. $SU(n)$ is the group of all n-dimensional unitary operators, A, such that $\det(A) = 1$.

Proposition 2.11. *We state these facts about $SO(n, \mathbb{R})$ from algebraic topology*

1. *$SO(n, \mathbb{R})$ is path-connected but not simply connected*
2. *There exists a double cover of $SO(n, \mathbb{R})$ called $Spin(n)$, and $Spin(n)$ is simply connected for $n > 2$.*
3. *$Spin(3) \cong SU(2)$*

By a double covering, we mean that there is a surjective group homomorphism $\rho : Spin(n) \to SO(n, \mathbb{R})$, where the fiber of each element in $SO(n, \mathbb{R})$ corresponds to two elements in $Spin(n)$. In the case of $n = 3$ we get $SU(2)/\{\pm \mathbb{I}\} = SO(3, \mathbb{R})$.

For any Lie group, G, and $y \in G$, define $L_y : G \to G$ by $L_y(g) = yg$. L_y and the inverse function $L_{y^{-1}}$ are C^∞ by definition, and thus L_y is a homeomorphism of G.

Proposition 2.12. *Let G be a connected Lie group and $U \subseteq G$ any neighbourhood of e, then $G = \langle U \rangle$.*

Proof. Let $H = \langle U \rangle$. For any $h \in H$, $L_h(U) \subseteq H$. Since L_h is a homeomorphism, $L_h(U)$ is an open neighbourhood of h and $L_h(U)$ is open in H. $H \subseteq \bigcup L_h(U)$ so H is open. Let $g \in G \setminus H$, and say $x \in L_g(U) \cap H$, then $x = gu \in H$ for some $u \in H$. By closure, $xu^{-1} = g \in H$, a contradiction, thus $L_g(U) \cap H = \emptyset$. Hence $G \setminus H = \bigcup L_g(U)$ is open, so H is closed. Since H is clopen and non-empty in a connected space, $H = G$. □

2.1.3 Lie Algebras

Definition 2.13. *A Lie algebra $\mathfrak{g}$ is a vector space and a skew-symmetric bilinear map, $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, which satisfies the Jacobi identity*

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0 \quad \textit{for all } X, Y, Z \in \mathfrak{g}$$

Definition 2.14. *A Lie algebra homomorphism is a map $\alpha : \mathfrak{g} \to \mathfrak{h}$ between two Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, satisfying $\alpha([X,Y]) = [\alpha(X), \alpha(Y)]$*

Definition 2.15. *Let G be a Lie group and let $\Psi : G \to \operatorname{Aut} G$ by $\Psi_g(X) = gXg^{-1}$. Define $\operatorname{Ad} : G \to \operatorname{Aut} T_eG$ by $\operatorname{Ad}(g) = (d\Psi_g)_e$. Define $\operatorname{ad} : T_eG \to \operatorname{End}(T_eG)$ by $\operatorname{ad}(X) = d(\operatorname{Ad})_X$.*

Note $\operatorname{Ad}(g)$ is itself a mapping on T_eG. Furthermore, $\operatorname{ad}(X)$ is a map from $T_x(G) \to T_{\operatorname{Ad}(X)}(\operatorname{Aut} T_eG)$, so that ad is a trilinear map.

For any $X, Y \in T_eG$ we may define $[X,Y] = \operatorname{ad}(X)(Y)$, so that each Lie group is naturally endowed with a Lie algebra on its tangent plane. From here on, if G is a Lie group, $\mathfrak{g}$ will be the Lie algebra naturally associated with G. We note that $\operatorname{ad}(X)(Y) = XY - YX$ if G is a matrix Lie group.

Proposition 2.16. *Let G be a Lie group, and $X \in \mathfrak{g}$, define the vector field V_X on G by $V_X(g) = (dL_g)_e(X)$ for $g \in G$. Let φ_X be the maximal integral curve of V_X with initial condition e. Then*

1. *$\varphi_X : \mathbb{R} \to G$*
2. *φ_X is a homomorphism*
3. *$\varphi'_X(t) = (dL_{\varphi(t)})_e(X)$*

Definition 2.17. *Let $\exp : \mathfrak{g} \to G$ by $\exp(X) = \varphi_X(1)$.*

Proposition 2.18. *The following properties hold for $\exp : \mathfrak{g} \to G$*

1. *$\varphi_{\lambda X}(t) = \varphi_X(\lambda t)$*
2. *$\exp(tX) = \varphi_X(t)$*
3. *$\exp((s+t)X) = \exp(sX) \circ \exp(tX)$*
4. *$T_0(\exp(X)) = X$*

Proof. (1) follows from the uniqueness of φ_X by the uniqueness of maximal integral curves. (2) follows by (1) and evaluating at $t = 1$. (3) follows from φ_X being a homomorphism. For (4), $T_0(\exp(X)) = (d/dt)(\exp(tX))\big|_0 = \varphi'_X(0) = X$. □

Theorem 2.19 (Lie Group-Lie Algebra Correspondence)**.**
Let $\rho : G \to H$ be a Lie group homomorphism. Then the following diagram commutes.

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{(d\rho)_e} & \mathfrak{h} \\ {\scriptstyle \exp}\downarrow & & \downarrow{\scriptstyle \exp} \\ G & \xrightarrow{\rho} & H \end{array}$$

Proof. Let $X \in \mathfrak{g}$, and let $\varphi(t) = \rho(\exp_G(tX))$ be a one-parameter subgroup of H. Then

$$\frac{d}{dt}\varphi(t)\big|_{t=0} = T_e(\rho)\ T_0(\exp_G(X)) = T_e(\rho)X = \varphi'(0)$$

Since $\varphi(t) = \exp(t\ \varphi'(0))$, this gives that $\varphi(1) = \exp_H(T_e(\varphi)X)$, and so $\exp_H \circ (d\rho)_e = \rho \circ \exp_G$. □

Since the differential of exp at 0 is e, the image of exp contains a neighbourhood of e. Thus if G is connected, G is completely determined by T_eG.

Theorem 2.20. *Let G and H be Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. Further suppose G is simply connected. A linear map $\alpha : \mathfrak{g} \to \mathfrak{h}$ is the differential of a Lie group homomorphism $A : G \to H$ if and only if α is a Lie algebra homomorphism.*

2.2 Representation Theory

2.2.1 Terminology

Definition 2.21. *Let G be a Lie group, V a finite-dimensional vector space, and $\rho : G \to GL(V)$ a Lie group homomorphism. Then (ρ, V) is called a Lie group representation.*

Definition 2.22. *Let $\mathfrak{g}$ be a Lie algebra, V a finite-dimensional vector space, and $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ a Lie group homomorphism. Then (ρ, V) is called a Lie group representation. By $\mathfrak{gl}(V)$ we mean the entire space generated by $GL(V)$ under $[\cdot, \cdot]$.*

When the homomorphism is clear, we often just write gv for $\rho(g)v$, we then simply refer to the vector space V as the representation; futhermore, a representation gives V the structure of a G-module. We also see the tensor product of two representations (the vector spaces) is a representation when we define $g(v \otimes w) = gv \otimes gw$.

Definition 2.23. *Let V be a representation of a group G, or an algebra $\mathfrak{g}$. A vector subspace W of V is a subrepresentation of V if it is invariant under the action of G, or $\mathfrak{g}$. V is irreducible if there are no proper nonzero G-invariant subspaces of V.*

Proposition 2.24. *Suppose G, or $\mathfrak{g}$, has no nonzero solvable ideals (semi-simple); then any representation of G, or $\mathfrak{g}$, is a direct sum of irreducible representations.*

Since representations of semi-simple objects are decomposable into irreducible representations, to study the representations of a semi-simple object we only need to consider its irreducibles and how an arbitrary representation decomposes.

Definition 2.25. *A representation on a complex vector space, V, is called unitary if the action of G, preserves the inner product on V (viewing V as a Hilbert space). That is*

$$\langle gv|gw\rangle = \langle v|w\rangle \quad \forall v, w \in V$$

Definition 2.26. *Let V be a real vector space, the complexification of V is $V_{\mathbb{C}} = V \otimes_{\mathbb{R}} \mathbb{C}$.*

The complexification of a vector space is the original vector space except that our scalars are now from $\mathbb{C}$ instead of $\mathbb{R}$, so $V_{\mathbb{C}} \cong V \oplus iV$.

Proposition 2.27. *Let $\mathfrak{g}$ be a finite-dimensional real Lie algebra. There is a unique extension of $[\cdot,\cdot]$ on $\mathfrak{g}$ to $\mathfrak{g}_{\mathbb{C}}$, making it into a complex Lie algebra. If $\mathfrak{h}$ is a complex Lie algebra, then every Lie algebra homomorphism of $\mathfrak{g}$ into $\mathfrak{h}$ extends uniquely to a (complex) Lie algebra homomorphism of $\mathfrak{g}_{\mathbb{C}}$ into $\mathfrak{h}$. $\mathfrak{h}$ is an irreducible representation of $\mathfrak{g}$ if and only if it is an irreducible representation of $\mathfrak{g}_{\mathbb{C}}$.*

Any endomorphism, X, of a complex vector space can be written uniquely as $X = X_s + X_n$ where X_s is diagonalizable and X_n is nilpotent, and X_s and X_n commute. This is called the *Jordan Decomposition*.

Proposition 2.28 (Preservation of Jordan Decomposition)**.**
Suppose $\mathfrak{g}$ has no nonzero solvable ideals. For any $X \in \mathfrak{g}$ there are $X_s, X_n \in \mathfrak{g}$ such that for any representation $\rho : \mathfrak{g} \to \mathfrak{gl}(V)$ we have

$$\rho(X)_s = \rho(X_s), \quad \rho(X)_n = \rho(X_n)$$

2.2.2 Irreducible Representations of $\mathfrak{sl}_2(\mathbb{C})$

We want to know all of the irreducible representations of $SU(2)$ in physics. So we will find the irreducible representations of the Lie algebra $\mathfrak{su}(2)$. By direct computation, we find that

$$\mathfrak{su}(2) = \left\{ \begin{pmatrix} ic & -a+bi \\ (a+bi) & -ic \end{pmatrix} : a, b, c \in \mathbb{R} \right\}$$

So even though $\mathfrak{su}(2)$ is the group of 2×2 complex traceless anti-hermitian matrices, it is a real Lie algebra, not a complex Lie algebra. Immediately we see

$$\begin{aligned} \mathfrak{su}(2)_{\mathbb{C}} &= \mathfrak{sl}_2(\mathbb{C}) \\ &= \left\{ \begin{pmatrix} \alpha & \beta \\ \gamma & -\alpha \end{pmatrix} : \alpha, \beta, \gamma \in \mathbb{C} \right\} \end{aligned}$$

We choose the basis for $\mathfrak{sl}_2(\mathbb{C})$ given by

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

and note that $[H, X] = 2X$, $[H, Y] = -2Y$, and $[X, Y] = H$.

If V is an irreducible representation of V, then by the preservation of Jordan decomposition we must have that the action of H on V is diagonalizable. This gives a decomposition of V into eigenspaces under the action of H, so $V = \bigoplus V_\alpha$ where $H(v) = \alpha v$ if $v \in V_\alpha$. We note that

$$H(X(v)) = X(H(v)) + [H, X](v) = X(\alpha v) + 2X(v) = (\alpha + 2)X(v)$$

so X maps $V_\alpha \to V_{\alpha+2}$ and similarly Y maps $V_\alpha \to V_{\alpha-2}$. Let α_0 be some eigenvalue that actually occurs in the decomposition then $\bigoplus_{\mathbb{Z}} V_{\alpha_0+2n}$ is $\mathfrak{sl}_2(\mathbb{C})$ invariant because it is invariant under the basis vectors. Since V was irreducible, $V = \bigoplus_{\mathbb{Z}} V_{\alpha_0+2n}$ and the eigenvalues of the action of H form an unbroken (finite) string $\beta, \beta+2, ..., \beta+2k = n$, we will denote V by $V^{(n)}$.

Take $v \in V^{(n)}$, and define $W = \mathbb{C}\{v, Y(v), Y^2(v), \dots\}$. By direct calculation we find that W is invariant under the action of X, Y, and H, and thus $W = V^{(n)}$, so V is completely determined by n. From these calculations, we also find that $X(Y^k(v)) = k(n - k + 1)(Y^{k-1}(v))$ and $dim(V_\alpha) = 1$ for each α. Since $V^{(n)}$ is finite dimensional, there is a minimal power, m, of Y such that $Y^m(v) = 0$. This tells us $0 = X(Y^m(v)) = m(n - m + 1)Y^{m-1}(v)$ so $n = m - 1$, and thus $n \in \mathbb{N}$.

As a result we have proven the following statement

Theorem 2.29 (The Irreducible Representations of $\mathfrak{sl}_2(\mathbb{C})$)**.**
For each $n \in \mathbb{N}$ the representation $V^{(n)}$ of $\mathfrak{sl}_2(\mathbb{C})$ is an $(n+1)$-dimensional irreducible representation, where H has eigenvalues $n, n-2, \dots, -n+2, -n$ when acting on $V^{(n)}$. These are all the irreducible representations. Furthermore $V^{(n)} = \mathrm{Sym}^n(V)$, where $V \cong \mathbb{C}^2$.

Proof. The only part that we haven't seen is that $V^{(n)} = \mathrm{Sym}^n(V)$ but this is clear when we compare the eigenvalues of H on $\mathrm{Sym}^n(V)$. □

By thinking about how eigenvalues of symmetric powers, wedge products, and tensor powers behave, we can decompose any representation into irreducibles. To illustrate this, we decompose $\mathrm{Sym}^2 V \otimes \mathrm{Sym}^3 V$ into irreducibles. The eigenvalues are $-2, 0, 2$ and $-3, -1, 1, 3$ respectively, and so the eigenvalues of the tensor product are the twelve sums $(-2) + (-3), (-2) + (-1), \dots, 2 + 3$. The eigenvalues ± 5 occur once, ± 3 occur with multiplicity 2, and ± 1 occur with multiplicity 3. So we see $\mathrm{Sym}^2 V \otimes \mathrm{Sym}^3 V \cong \mathrm{Sym}^5 V \oplus \mathrm{Sym}^3 V \oplus V$.

3 Mathematics of Quantum Mechanics

Our goal in this section is to give a rigorous formulation of quantum mechanics. We then look to the role played by symmetries in quantum mechanics with the introduction of Wigner's theorem.

3.1 The Basics

3.1.1 Quantum Foundations

Here we will follow the work of Mackey as in [9] and [10], and try to develop quantum mechanics by viewing classical mechanics and quantum mechanics as cases of a more general framework. Our notation will follow that of [14].

In classical mechanics the phase space of a dynamical system in 3-dimensions of N particles is a $6N$-dimensional Borel subspace, Ω, of $\mathbb{R}^{6N}$. A point $\omega \in \Omega$ carries 6 pieces of information for each of the N particles. Namely, 3 Cartesian positions and 3 Cartesian momenta, and is called a state. For each $i \in \{1, \ldots, N\}$ and $j \in \{x, y, z\}$ let q_{ij}, p_{ij} denote the j-th component of the i-th particles position (q) and momentum (p) respectively. The development of our state ω in phase space is then governed by Hamilton's equations:

$$\frac{dq_{ij}}{dt} = \frac{\partial H}{\partial p_{ij}} \qquad\qquad \frac{dp_{ij}}{dt} = -\frac{\partial H}{\partial q_{ij}}$$

Where H is a function which happens to give the equations of motion and physicists have access to, typically corresponding to the total energy of the system. Most importantly, for each pair (i, j) these are two coupled first order ODEs, thus we know for any initial condition ω_0 at t_0 that we may compute $\omega(t + t_0)$ for any t. Hence, understanding the trajectories of solutions to these ODEs in classical phase space is the same as fully understanding the system.

For physical reasons no correct description of quantum phenomena can permit us to know both the position and momentum of a particle in phase space simultaneously. The measurements must interfere. However, sampling of identical statistical ensembles can allow us to form statistical descriptions of the system which will not interfere. This means that we will have to frame our description of quantum mechanics so that what we measure is fundamental, and that statistical descriptions will be necessary.

Let α be a probability measure on Ω (i.e., $\alpha(\Omega) = 1$), then we will call α a *statistical state*, and $\alpha(E)$ is the probability that the true configuration of the dynamical system is in the subset E of phase space. An *observable* will be a (real) function on

Ω, which will correspond to some measured quantity of physical interest; momentum, energy, etc. Say α is a statistical state, I is a Borel subset of $\mathbb{R}$, and f is an observable, we see that $\alpha(f^{-1}(I))$ is the probability that a measurement of f will lead to a value in I. By a measurement of f we mean $f(\omega)$ for some ω.

The problem is that these definitions of statistical state and observable do not translate nicely to the phase space of quantum mechanics. The quantum analog of Ω is an extremely messy thing. However, if we generalize these concepts to an orthocomplemented lattice we will find that we can avoid working in quantum phase space all together.

Let $\mathscr{L}$ be a nonempty set. $\mathscr{L}$ is partially ordered if there is a binary relation, $\leq$, on $\mathscr{L}$ which is reflexive, anti-symmetric, and transitive. If A is a nonempty subset of $\mathscr{L}$, we will denote the (unique) supremum of A by $\bigvee_{a\in A} a$, similarly $\bigwedge_{a\in A} a$ will denote the (unique) infimum of A. If $A = \mathscr{L}$, the infimum is denoted 0, and the supremum is denoted 1. If infimums and supremums exist for every finite subset of $\mathscr{L}$, and 0 and 1 elements exist and are distinct, then $\mathscr{L}$ is a *bounded lattice*.

Definition 3.1. *An orthocomplementation on a bounded lattice $\mathscr{L}$ is a bijection on $\mathscr{L}$, $\perp$: $a \to a^{\perp}$, satisfying the following for every $a, b \in \mathscr{L}$*

1. $a^{\perp\perp} = a$
2. *If $a \leq b$, then $b^{\perp} \leq a^{\perp}$*
3. $a \vee a^{\perp} = 1$ *and* $a \wedge a^{\perp} = 0$

We will write $a \perp b$ if $a \leq b^{\perp}$. We also note that a Boolean algebra is an orthocomplemented lattice with the distributive property on $\vee$ and $\wedge$.

Definition 3.2. *Let $\mathscr{L}$ be an orthocomplemented lattice. A probability measure on $\mathscr{L}$ is a function $\mu : \mathscr{L} \to [0, 1]$ satisfying*

1. $\mu(0) = 0$ *and* $\mu(1) = 1$
2. *If $\{a_i\}$ is a countable collection of members of $\mathscr{L}$ and $a_i \perp a_j$ for each $i \neq j$, then $\mu(\bigvee a_i) = \sum \mu(a_i)$*

Definition 3.3. *Let $\mathscr{B}$ denote all Borel subsets of $\mathbb{R}$ and let $\mathscr{L}$ be an orthocomplemented lattice. An $\mathscr{L}$-valued measure is a function $\mathscr{M} : \mathscr{B} \to \mathscr{L}$ mapping $E \to L_E$ satisfying:*

1. $E \cap F = \emptyset$ *implies* $L_E \perp L_F$
2. $L_{E_1 \cup E_2 \cup \ldots} = L_{E_1} \vee L_{E_2} \vee \ldots$
3. $L_{\emptyset} = 0$ *and* $L_{\mathbb{R}} = 1$

These definitions illustrate how we should re-define statistical states and observables: *statistical states* (or just *states*) are defined to be probability measures on $\mathscr{L}$; and *observables* are defined to be $\mathscr{L}$-valued measures.

Define $\mathscr{L}_\Omega$ to be the orthocomplemented lattice of all Borel subsets of the classical phase space Ω, equipped with the partial order given by set-theoretic containment $\subseteq$, and orthocomplement mapping $E \to \Omega \backslash E$. In this case, $E \perp F$ means E and F are disjoint subsets of phase space, and $\vee$ becomes set-theoretic unions.

In quantum mechanics our equivalent to $\mathscr{L}_\Omega$ will be $\mathscr{L}_\mathcal{H}$, the orthocomplemented lattice of all projection operators of a separable complex Hilbert space $\mathcal{H}$. The partial order will be given as before on subsets of $\mathcal{H}$, and complementation takes the subspace $A \subseteq \mathcal{H}$ to $A^\perp$ defined by the inner product on $\mathcal{H}$, and so on.

One question is why do we use the orthocomplemented lattice we've chosen? The short answer is that this choice works extremely well in describing physical phenomenon (in that no issues have been found). There are lengthy arguments about what one may expect from the spaces that describe quantum systems, but still none completely rule out, on non-aesthetic grounds, the possibility of real or quaternionic Hilbert spaces, or something even more pathological.

Before investigating $\mathscr{L}_\mathcal{H}$ we take a look at a final concept about our states

Definition 3.4. *Let $\{\alpha_i\}$ be a collection of states of an arbitrary lattice $\mathscr{L}$, not all identical. The probability measure $\alpha = \sum \gamma_i \alpha_i$ where $\sum \gamma_i = 1$ and $\gamma_i \in [0,1]$ is called a mixed state. Any state which cannot be written as a (non-trivial) convex combination of others is called a pure state.*

3.1.2 The Vector-Operator Language

From here out, we will look at the properties of $\mathscr{L} = \mathscr{L}_\mathcal{H}$. Let $\psi \in \mathcal{H}$ be a unit vector, for each projection, P, ψ defines a probability measure on $\mathscr{L}_\mathcal{H}$ by $\alpha_\psi(P) = \langle P\psi | \psi \rangle$. Furthermore, we have that convex combinations of $\{\alpha_i\}$ will be a state. The converse is also true, that is

Theorem 3.5. *Every probability measure on $\mathscr{L}_\mathcal{H}$ is of the form $\sum \gamma_i \alpha_{\psi_i}$, where $\sum \gamma_i = 1$ and $\gamma_i \in [0,1]$. Furthermore, the $\{\psi_i\}$ can be chosen to be a complete orthonormal set of $\mathcal{H}$.*

A proof of the above theorem can be found in [4] and played a major part in formulating the fundamentals of quantum mechanics.

Definition 3.6. *Let $\{\psi_i\}$ be a complete orthonormal set in $\mathcal{H}$ and $\alpha = \sum \gamma_i \psi_i$ a state (mixed or pure). Define the von Neumann density operator, A, corresponding to α to be the unique bounded linear operator such that $A(\psi_i) = \gamma_i \psi_i$, for all ψ_i.*

The existence and uniqueness of such an operator is assured by the Riesz Representation Theorem. We then have for any state $\alpha = \sum \gamma_i \alpha_{\psi_i}(P)$, that

$$\alpha = \sum \gamma_i \alpha_{\psi_i}(P) = \sum \gamma_i \langle P\psi_i | \psi_i \rangle = \mathrm{Tr}(PA)$$

Hence we find a one-to-one correspondence between statistical states, α, and self-adjoint operators with unit trace and non-negative eigenvalues, PA.

Furthermore, we see the state represented by A is pure if and only if A has a one-dimensional range. If ψ is a unit vector in this range $\mathrm{Tr}(PA) = \langle P\psi|\psi\rangle$, and hence pure states are those of the form α_ψ. Thus ψ and $e^{ia}\psi$ define the same state for $a \in \mathbb{R}$. That is, the pure states correspond to the unit rays of the Hilbert space, which we will look at in further detail in Wigner's Theorem.

We recall that the quantum mechanical observables are the $\mathscr{L}_\mathcal{H}$-valued measures on $\mathbb{R}$, and in terms of our definitions these are the projection valued measures. From the spectral theorem for self-adjoint operators we know then see that there is a one-to-one correspondence between observables and the self-adjoint operators in $\mathcal{H}$.

We summarize our developments in the following theorem

Theorem 3.7 (von Neumann's Unification Theorem)**.**
Let $\mathcal{H}$ be a separable infinite-dimensional Hilbert space. Then for any quantum mechanical system described by the lattice $\mathscr{L}_\mathcal{H}$

- *Pure states correspond to unit rays of $\mathcal{H}$*
- *Observables correspond one-to-one with self-adjoint operators on $\mathcal{H}$*

Let A be the self-adjoint operator corresponding to an observable, and ψ be a unit-vector corresponding to the state of the system. Let P_E^A be the projection valued measure associated with A via the spectral theorem. Then $\left\langle P_E^A\psi \middle| \psi \right\rangle$ is the result of measuring this observable.

We often write $\left\langle P_E^A\psi \middle| \psi \right\rangle$ as $\left\langle \psi \middle| P_E^A\psi \right\rangle$ or $\langle\psi| P_E^A |\psi\rangle$.

An observable A has a *pure point spectrum* if there is an orthonormal basis $\{\psi_i\}$ such that $A(\psi_i) = \lambda_i\psi_i$. If any λ_i occurs with multiplicity greater than 1, then that eigenvalue is called *degenerate*. The eigenvectors for the operator A will be called *eigenstates*.

Since we know self-adjoint operators with no non-degenerate eigenvalues have orthogonal eigenvectors we get the very useful computational corollary

Corollary 3.8. *Suppose A is an observable with pure point spectrum and no degenerate eigenvalues. Further suppose that $\{\psi_i\}$ is the orthonormal eigenbasis for A. Suppose the system is in the state ψ. Then $\psi = \sum \gamma_i\psi_i$ where $\gamma_i = \langle\psi|\psi_i\rangle$. Furthermore, $\sum|\langle\psi|\psi_i\rangle|^2 = 1$ and the probability that the observable A has value λ_i is $|\langle\psi|\psi_i\rangle|^2$, and the probability it has a value other than one of the λ_i is 0.*

We see from the spectral theorem that the expectation value of the observable A for any state ψ is given by $\langle \psi | A\psi \rangle$.

In the case of an infinite-dimensional $\mathcal{H}$, we identify the vectors ψ with a function $\psi \in L^2(\mathbb{R}^n)$

$$\langle \psi | \psi \rangle = \int_{R^n} \psi(\mathbf{x})^* \psi(\mathbf{x}) d\mathbf{x} = 1$$

And so the probability that a particle in n-spatial dimensions is in a Borel subset $E \subseteq \mathbb{R}^n$ is

$$\langle \psi | \psi \rangle = \int_{E} \psi(\mathbf{x})^* \psi(\mathbf{x}) d\mathbf{x}$$

From this, we clearly see how if ψ_1 and ψ_2 are two vectors corresponding to the same ray in Hilbert space then their inner products are the same. Hence operations on rays will be well defined.

From here on, when we talk about a state we will always talk about the vectors in Hilbert space to which it corresponds. Similarly, when we speak of observables, we will refer to the self-adjoint operators to which they correspond. When we speak of a vector in $\Psi \in \mathcal{H}$ we may denote it as $|\Psi\rangle$ so its natural dual will be $\langle \Psi |$.

Example 3.9. *Let $\Psi : \mathbb{R}^3 \to \mathbb{C}$ be a quantum state. Define $\hat{\mathbf{x}}$ to be the operator such that $\hat{\mathbf{x}}\Psi(\mathbf{x}) = \mathbf{x}\Psi(\mathbf{x})$. Then $\hat{\mathbf{x}}$ does NOT have a pure point spectrum. Let $|\mathbf{x}\rangle$ be the eigenstate of $\hat{\mathbf{x}}$ with eigenvalue $\mathbf{x}$. Then we have $\langle \mathbf{x}' | \mathbf{x} \rangle = \delta(\mathbf{x}' - \mathbf{x})$ where δ is the Dirac delta. Furthermore, $\langle \mathbf{x} | \Psi \rangle = \Psi(\mathbf{x})$.*

3.1.3 Time-Dependent States

Up until now we have only considered states that are not time-dependent, so called *stationary states* of the system. However, adding time-evolution is straight-forward. If Ψ is a state of the system at any time, then t units of time later the state should be represented by the unit vector $U_t(\Psi)$, where U_t is a unitary operator with inverse $U_t^{-1} = U_{-t}$. This defines a continuous one-parameter unitary group, which we call the *dynamical group* of the system. Furthermore, we may apply Stone's theorem and write $U_t = e^{-itH}$ for some self-adjoint operator H, which we will call our (quantum) Hamiltonian. The eigenvalues of the operator H are the possible energy values of the system.

Suppose $\Psi(0)$ is the state of the system at time $t = 0$, then $\Psi(t) = U_t(\Psi(0)) = e^{-itH}\Psi(0)$. Taking a derivative we get

$$\begin{aligned}\frac{d}{dt}\Psi(t) &= \frac{d}{dt}(e^{-itH}\Psi(0)) \\ &= -iHe^{-itH}\Psi(0) \\ &= -iH(\Psi(t))\end{aligned}$$

This differential equation which served as the original basis for quantum mechanics is known as *Schrödinger's Equation.*

$$\frac{d}{dt}\Psi(t) = -iH(\Psi(t))$$

Typically a physicist will create the appropriate H corresponding to a physical system and then solve the resultant PDE by separation of variables.

3.2 Symmetries and Wigner's Theorem

In the following section we look at the result known as Wigner's Theorem, originally produced by Wigner in [16], and formalized by numerous authors since. In particular, we follow the proof by Weinberg in [15]. For a proof that removes all physics, see [1].

We are interested in symmetries in quantum mechanics because of ways they can simplify problems or physical intuition they can provide. For example, Noether's Theorem[1] states any system invariant under a continuous symmetry will have a quantity associated with that symmetry which is conserved. Invariance under time translation results in conservation of energy, spatial translations leads to conservation of momentum, and the fact that the scalar and vector potentials associated with electromagnetism do not depend on their zero points, only field differences, leads to charge conservation [11]. In particular, if G is a Lie group describing a symmetry of our system, then it is encoded by (projective) unitary representations in the observables.

There is no way in general to predict the symmetries that will occur in an undiscovered phenomenon. The fundamental Lagrangians (a Legendre transform of our Hamiltonian) that describe physical phenomena are mostly written by "feel" with foresight based on experiment and known recipes that usually work in outlining phenomena. For example, there is no reason that we could predict that the states describing the quarks of chromodynamics are invariant under the action of the $SU(3)$ Lie group. But once we do find this, we know there is an associated "colour-conservation" law.

[1] See [13] for a rigorous mathematical statement.

In this section, bold letters will correspond to rays, and non-bold letters correspond to vectors. $\mathcal{H}$ and $\mathcal{H}'$ are Hilbert spaces, and $\mathcal{R}(\mathcal{H})$ is the ray space associated with the Hilbert space $\mathcal{H}$, and similarly with primes.

Definition 3.10. *Let $\mathbf{\Psi}$ and $\mathbf{\Phi}$ be rays corresponding to vectors Ψ and Φ respectively in a Hilbert space. Define an inner product, called the ray product, between two rays by*

$$\mathbf{\Psi} \cdot \mathbf{\Phi} := |\langle \Psi | \Phi \rangle|$$

If a system is invariant under some transformation then the outcomes of measurements on the untransformed system should be the same as the transformed system, formally this means that the ray product of transformed rays will be the same as the ray product of the untransformed rays.

Definition 3.11. *Let $U : \mathcal{H} \to \mathcal{H}'$ and $T : \mathcal{R}(\mathcal{H}) \to \mathcal{R}(\mathcal{H}')$ a transformation of their respective ray spaces. We say U is compatible with T if $U\Psi \in T\mathbf{\Psi}$ for all $\Psi \in \mathcal{H}$.*

We note that if U is a unitary or anti-unitary transformation of Hilbert spaces that the transformation $T : \mathbf{\Psi} \to \mathbf{U\Psi}$ induced by U is compatible with U and preserves inner products. Wigner's theorem states that no other inner product preserving mappings of ray space exist.

Theorem 3.12 (Wigner's Theorem)**.**
Let T be a mapping of the unit ray space of $\mathcal{H}$ into the unit ray space of $\mathcal{H}'$ and suppose for every $\Phi, \Psi \in \mathcal{H}$ that

$$T\mathbf{\Phi} \cdot T\mathbf{\Psi} = \mathbf{\Phi} \cdot \mathbf{\Psi}$$

then there exists a mapping $U : \mathcal{H} \to \mathcal{H}'$ compatible with T, which is unitary or anti-unitary.

Proof. Let T be as given. If $\mathbf{\Psi}$ is any given ray, we will use Ψ as a vector representative from the ray. For any ray $\mathbf{\Psi}$ let $\mathbf{\Psi}' = T\mathbf{\Psi}$ so that $\Psi' \in \mathbf{\Psi}'$.

Let $\{\Psi_k\}$ be a set of complete orthonormal state vectors[2] in $\mathcal{H}$. By hypothesis

$$|\langle \Psi'_k | \Psi'_l \rangle|^2 = |\langle \Psi_k | \Psi_l \rangle|^2 \tag{1}$$

where $\langle \Psi_k | \Psi_l \rangle = \delta_{kl}$ in turn, implying the $\{\Psi'_k\}$ are also orthonormal and and complete (else we can show $\{\Psi_k\}$ is not complete).

[2]If $\dim \mathcal{H} = 1, 2$, or 3, our proof terminates in straightforward spots.

Single out Ψ_1 and for $k \neq 1$ define

$$\Upsilon_k = \frac{1}{\sqrt{2}}[\Psi_1 + \Psi_k] \in \boldsymbol{\Upsilon_k}$$

Any state vector $\Upsilon'_k \in \boldsymbol{\Upsilon'_k}$ may be expanded as

$$\Upsilon'_k = \sum c_{kl} \Psi'_l$$

and by (1) we have $|c_{kk}| = |c_{k1}| = 1/\sqrt{2}$, and $c_{kl} = 0$ for $l \neq k$ and $l \neq 1$. With hindsight, we can choose our ray representatives Υ'_k and Ψ'_k such that $c_{kk} = c_{k1} = 1/\sqrt{2}$ with no phase factor, which we will denote $U\Upsilon_k$ and $U\Psi_k$. Giving

$$U\frac{1}{\sqrt{2}}[\Psi_1 + \Psi_k] = U\Upsilon_k = \frac{1}{\sqrt{2}}[U\Psi_1 + U\Psi_k]$$

defining U for some vectors, but not a general vector.

We can extend the mapping T from unit ray space to all rays by $T(\boldsymbol{\Psi}) = \|\Psi\| T(\hat{\boldsymbol{\Psi}})$ where $\hat{\boldsymbol{\Psi}}$ is the unique unit ray attained by scaling down the vectors of $\boldsymbol{\Psi}$.

Let $\Psi \in \boldsymbol{\Psi}$, not necessarily a unit ray, and $\Psi' \in \boldsymbol{\Psi'}$, then we have

$$\Psi = \sum C_k \Psi_k \qquad \text{and} \qquad \Psi' = \sum C'_k U\Psi_k$$

From (1) on $|\langle \Psi_k | \Psi \rangle|^2 = |\langle U\Psi_k | \Psi' \rangle|^2$ we get $|C_k|^2 = |C'_k|^2$, similar expansions and products with Υ_k and Υ'_k gives (for $k \neq 1$) $|C_k + C_1|^2 = |C'_k + C'_1|^2$. Taking the ratios of these quantities and looking at the real and imaginary parts we eventually realize

$$\frac{C_k}{C_1} = \frac{C'_k}{C'_1}$$

or else

$$\frac{C_k}{C_1} = \left(\frac{C'_k}{C'_1}\right)^*$$

And it can be shown, by contradiction, that the same choice of non-conjugate or conjugate must be chosen for all k by considering the same process on a state defined as a linear combination of 3 vectors. We may thus define $U\Psi \in \boldsymbol{\Psi'}$ with phase chosen so that $C_1 = C'_1$ or $C_1 = C'^*_1$ and so

$$U\left(\sum C_k \Psi_k\right) = \sum C_k U\Psi_k \tag{2}$$

or else

$$U\left(\sum C_k \Psi_k\right) = \sum C_k{}^* U\Psi_k \tag{3}$$

It can be shown, again by contradiction with the real and imaginary coefficients that arise, that for any fixed symmetry transformation T, that the same choice of defining U by (2) or (3) must be made for all sets of C_k. Hence, all state vectors satisfy either (2) or (3), and so U (which is now defined for an arbitrary vector) is seen to be unitary or anti-unitary respectively. □

Example 3.13. *Let $\hat{\mathbf{x}}$ and $|\mathbf{x}\rangle$ as in Ex. 3.9. Define the parity or mirror-symmetry operator P by $P|\mathbf{x}\rangle = |-\mathbf{x}\rangle$. We have $P^2 = \mathbf{1}$ and also $\langle \mathbf{x}'|P\mathbf{x}\rangle = \langle \mathbf{x}'|-\mathbf{x}\rangle = \delta(\mathbf{x}'+\mathbf{x})$ and $\langle P\mathbf{x}'|\mathbf{x}\rangle = \langle \mathbf{x}|-\mathbf{x}'\rangle^* = \delta(\mathbf{x}'+\mathbf{x})$ so P is self-adjoint. Thus $P^{-1} = P = P^\dagger$ and P is unitary. Thus P induces a symmetry transformation on the our quantum states, preserving probabilities. This corresponds to the fact that an experiment in a left-handed coordinate system or mirror universe should give all the same results as our own*[3].

By Wigner's theorem, if we have a quantum system which has a Lie group of symmetries, then the state space naturally carries a unitary representation of that symmetry group (defined up to rays). By the Lie group-algebra correspondence, we have an induced representation on the corresponding unitary Lie algebra. Since the Lie algebra of a unitary group is comprised of skew-hermitian matrices, there is a natural set of hermitian matrices (given by multiplication by i) which comprises the observables associated with the symmetry. So if X is an observable, iX is skew-hermitian, and so $exp(iX)$ is unitary, giving a nice relation to rotations and illustrating how the generators of symmetries are the related observables. We summarize this in the following remarkable theorem

Theorem 3.14. *Suppose a quantum system, $\mathcal{H}$, has a Lie group of symmetries, G. Then the Hilbert space of states carries a natural (projective) unitary representation of G. Furthermore, elements of the associated Lie algebra to G give the observables for the system.*

[3]Although, this turns out to be a complete lie when the electroweak interaction is involved. Both mirror symmetries and time-reversal $t \to -t$ turn out not to be symmetry transformations in experiment [18].

4 Composite Systems

This final section looks at how we can model things in quantum mechanics using representation theory and some consequences. For more, see [17] and [6].

4.1 Multi-Particle Systems

A quantum mechanical system often needs many Hilbert spaces to provide all of the relevant information to describe a particle in the system. This requires us to write the state of the system as a tensor product of states from each of the Hilbert spaces.

Suppose you have two quantum systems, $\mathcal{H}_1$ and $\mathcal{H}_2$, with sets of operators $\mathcal{O}_1$ and $\mathcal{O}_2$ respectively. Then the composite system is given by

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$$

This generalizes to composite systems with n components.

If the system is non-interacting then the operators on $\mathcal{H}$ take the form

$$A_1 \otimes \mathbb{I} + \mathbb{I} \otimes A_2$$

where $A_1 \in \mathcal{O}_1$ and $A_2 \in \mathcal{O}_2$. If the system is interacting, then the operators are free to take their most general form.

Let $\mathcal{H}$ be the state space for a single particle, then n identical[4] particles will have a state space given by $\mathcal{H}^{\otimes n}$. There is a natural action of $\mathfrak{S}_n$ on $\mathcal{H}^{\otimes n}$ given by permuting elements of $\mathcal{H}^{\otimes n}$. That is, if $\sigma \in \mathfrak{S}_n$, then

$$\sigma \cdot (\psi_1 \otimes \psi_2 \otimes \cdots \otimes \psi_n) = \psi_{\sigma(1)} \otimes \psi_{\sigma(2)} \otimes \cdots \otimes \psi_{\sigma(n)}$$

So $\mathcal{H}^{\otimes n}$ is seen to be a representation of $\mathfrak{S}_n$.

Definition 4.1. *Let $\psi \in \mathcal{H}^{\otimes n}$ be a state vector for a quantum system of n identical particles. If $\psi \in \operatorname{Sym}^n \mathcal{H}$, then ψ is called bosonic. If $\psi \in \bigwedge^n \mathcal{H}$, then ψ is called fermionic.*

We take the following statement to be an axiom of quantum mechanics: *All physical states occur as one dimensional representations of $\mathfrak{S}_n$. They are either all bosonic (composed of bosons) or fermionic (composed of fermions).* With this in mind, we find the action of $\mathfrak{S}_n$ on bosons and fermions is straightforward.

[4] We note that when we say the particles are identical we mean they are completely indistinguishable. Unlike the macroscopic world we cannot paint one particle red and one particle blue to keep track. This is not a problem with measurement, we simply require that identical particles remain purely indistinguishable, and suppose that nature will not give one particle special treatment by painting it blue when all the others are red.

Proposition 4.2. *Let ψ be a physical state for a quantum system of n identical particles, and $\sigma \in \mathfrak{S}_n$.*

1. *If ψ is bosonic, $\sigma \cdot \psi = \psi$.*
2. *If ψ is fermionic, $\sigma \cdot \psi = \operatorname{sgn}(\sigma)\psi$*

This leads us to the well-known *Pauli Exclusion Principle* as a corollary

Corollary 4.3. *No two identical fermions may occupy the same quantum state in the same quantum system simultaneously.*

Proof. Suppose the two particles are in the same quantum state ψ, so that the state of the two fermions in the combined system is $\psi \otimes \psi$. Since the particles are fermions, we have $\psi \otimes \psi \in \bigwedge^2 \mathcal{H}$, so $\psi \otimes \psi = 0$. This is not a valid quantum state. □

We now return back to arbitrary (non-identical) tensor products of states.

Definition 4.4. *Let $\{\mathcal{H}_\rangle\}$ be state spaces, and $\mathcal{H} = \bigotimes \mathcal{H}_i$ the composite state space. The indecomposable vectors in $\mathcal{H}$ are called entangled states.*

An important example of non-interacting spaces occurs in the description of an electron bound to a hydrogen atom. We write it as a tensor product of states from an infinite Hilbert space ($L^2(\mathbb{R}^3)$) and 3 finite-dimensional Hilbert spaces. The infinite dimensional Hilbert space is spanned by a complete orthonormal set of eigenfunctions, $\{|n\rangle\}$, for the Hamiltonian operator, so we will call it $\mathcal{H}_n$. Similarly, the 3 finite-dimensional spaces are spanned by complete orthonormal sets $\{|l\rangle\}$, $\{|m\rangle\}$, $\{|s\rangle\}$ (which have canonical choices for their elements), we denote these finite-dimensional spaces $\mathcal{H}_l$, $\mathcal{H}_m$, and $\mathcal{H}_s$. The information taken from $\mathcal{H}_n, \mathcal{H}_l$, and $\mathcal{H}_m$ are related to a particle's position and momentum, while the information, called *spin*, from $\mathcal{H}_s$ is a fundamental quantity that cannot be explained without relativistic quantum theory. In the case of the electron, $\mathcal{H}_s = \mathbb{C}^2$.

4.2 $SU(2)$ and Spin

Every classical object in 3-dimensional space has an angular momentum, $\vec{J} = \vec{r} \times \vec{p}$. This quantity depends on the choice of origin of our system, but is always conserved. For any classical object orbiting an axis, but not spinning, we may pretend all of its mass is concentrated at its center of mass, and then define $\vec{L}$ to be the angular momentum of this compacted particle. An example is the Earth orbiting the Sun while ignoring the rotation of Earth. For a classical object not orbiting any external axis, but spinning about an axis through its center of mass, we may compute just

this spinning part of the angular momentum $\vec{S}$. When we return to the most general system, we find that $\vec{J} = \vec{L} + \vec{S}$, so that $\vec{J}$ can be decomposed into $\vec{L}$ which describes the way it orbits, and $\vec{S}$ which describes the way it spins.

The name spin in quantum mechanics comes from the notion that the quantity from a particles $\mathcal{H}_s$ space behaves similar to the angular momentum of an object spinning about its axis. While the quantity from $\mathcal{H}_l$ behaves similar to the orbital angular momentum.

Viewing an electron as an object orbiting the nucleus, we expect both it's orbital and spin angular momenta to have $SO(3,\mathbb{R})$ symmetries, so that $\mathcal{H}_l$ and $\mathcal{H}_s$ carry representations of $SO(3,\mathbb{R})$. Thus turns out to be only the case for $\mathcal{H}_l$. As it turns out, $\mathcal{H}_s$ actually carries $SU(2)$ as a symmetry group. We note (by an index change) that the irreducible representations of $SU(2)$ can be characterized by numbers $s = 0, 1/2, 1, 3/2, \ldots$, with dimension $2s+1$, since representations of $SU(2)$ correspond to representations of $\mathfrak{su}(2)$ which correspond to representations of $\mathfrak{sl}_2(\mathbb{C})$. A particle whose $\mathcal{H}_s$ space is irreducible representation V^{2s} of $SU(2)$ is said to have spin s.

As mentioned, the existence of the $\mathcal{H}_s$ space cannot be explained without relativistic quantum theory, so we must take its existence as an axiom. Relativistic quantum theory in 3-dimensions also gives us the *spin-statistics theorem* (if we assume all particles have a positive energy) which we must also take as an axiom: *A particle with integral spin is a boson. A particle with half-integral spin is a fermion.*

Examples of bosons are the Higgs (spin-0), the photon (spin-1), and the graviton (spin-2). Fermions are the particles that make up matter, like electrons, quarks, protons, and neutrons (spin-1/2) called *spinors*. There are no fundamental spin-3/2 particles that have been observed, and their discovery would validate over three decades of supersymmetric quantum field theory research.

4.3 Spin Systems

For this discussion, we assume we have particles where the information presented by the spatial state-spaces is unimportant, and that only the information provided by their spin state is needed to describe the particle.

Suppose then that we have two particles with spins s_1 and s_2. Then we know that the composite system has spin state space

$$\mathcal{H} = V^{2s_1} \otimes V^{2s_2}$$

Since the tensor product of representations is a representation, we have that $\mathcal{H}$ is decomposable into irreducible representations. Considering the eigenvalues of the

irreducible representations of $\mathfrak{su}(2)$ we see that $\mathcal{H}$ is decomposable as

$$\mathcal{H} \cong V^{2s_2+2s_1} \oplus V^{2s_2+2s_1-2} \oplus \cdots \oplus V^{|2s_2-2s_1|}$$

Viewing spin as angular momentum this is sensible since we do not expect the total angular momentum to be greater in magnitude than the sum of the individual momenta (imagine the momentum vectors to be parallel), and lesser than their difference (imagine vectors anti-parallel).

Consider the specific example of a quantum system composed of two electrons. In this case,

$$\begin{aligned} \mathcal{H} &= V^1 \otimes V^1 \\ &\cong V^2 \oplus V^0 \end{aligned}$$

We choose the seemingly unnatural basis for $\mathcal{H}$ given by

$$\binom{1}{0} \otimes \binom{1}{0}, \quad \frac{1}{\sqrt{2}}\left(\binom{1}{0} \otimes \binom{0}{1} + \binom{0}{1} \otimes \binom{1}{0}\right), \quad \binom{0}{1} \otimes \binom{0}{1}$$

and

$$\frac{1}{\sqrt{2}}\left(\binom{1}{0} \otimes \binom{0}{1} - \binom{0}{1} \otimes \binom{1}{0}\right)$$

The first 3 are referred to as the triplet states, and the last vector is called the singlet state. Applying an arbitrary $SU(2)$ element to the singlet state, we see that the singlet state is $SU(2)$ invariant. This makes the isomorphism of $\mathcal{H}$ to $V^2 \oplus V^0$ explicit, and gives us a basis to decompose an arbitrary two-electron system into spin-1 and spin-0 parts, via the triplet and singlet states respectively.

From just this simple two electron system we can conclude with one of the most puzzling philosophical issues in quantum mechanics: quantum entanglement.

Proposition 4.5. *Let $\mathcal{H}$ be a 2-dimensional state space and $|\phi_a\rangle, |\phi_b\rangle$ an orthonormal basis. The two particle state*

$$|\Psi\rangle = \alpha\,|\phi_a\rangle \otimes |\phi_b\rangle + \beta\,|\phi_b\rangle \otimes |\phi_a\rangle$$

is entangled for all non-zero $\alpha, \beta \in \mathbb{C}$.

Proof. Suppose otherwise, that there exists $|\psi_1\rangle, |\psi_2\rangle \in \mathcal{H}$ such that $|\Psi\rangle = |\psi_1\rangle \otimes$

$|\psi_2\rangle$. Then

$$\begin{aligned}
|\Psi\rangle &= |\psi_1\rangle \otimes |\psi_2\rangle \\
&= (\langle\phi_a|\psi_1\rangle\,|\phi_a\rangle + \langle\phi_b|\psi_1\rangle\,|\phi_b\rangle) \otimes (\langle\phi_a|\psi_2\rangle\,|\phi_a\rangle + \langle\phi_b|\psi_2\rangle\,|\phi_b\rangle) \\
&= \quad \langle\phi_a|\psi_1\rangle\,\langle\phi_a|\psi_2\rangle\,(|\phi_a\rangle \otimes |\phi_a\rangle) \\
&\quad + \langle\phi_a|\psi_1\rangle\,\langle\phi_b|\psi_2\rangle\,(|\phi_a\rangle \otimes |\phi_b\rangle) \\
&\quad + \langle\phi_b|\psi_1\rangle\,\langle\phi_a|\psi_2\rangle\,(|\phi_b\rangle \otimes |\phi_a\rangle) \\
&\quad + \langle\phi_b|\psi_1\rangle\,\langle\phi_b|\psi_2\rangle\,(|\phi_b\rangle \otimes |\phi_b\rangle)
\end{aligned}$$

This implies $\langle\phi_a|\psi_1\rangle\,\langle\phi_a|\psi_2\rangle = 0$, so $\langle\phi_a|\psi_1\rangle = 0$ and $\alpha = 0$, or $\langle\phi_a|\psi_2\rangle = 0$ and $\beta = 0$. A contradiction in either case. □

From this we see there are two-particle states, like the singlet state, that cannot be described as two single-particle states: we cannot discuss one particle while ignoring the other. This is troubling for the following reason, suppose we construct a machine that measures the observable corresponding to

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

for a single particle system, which has eigenvalues ± 1. With two such machines, we can measure the quantity corresponding to the two particle operator $\sigma_z \otimes \mathbb{I} + \mathbb{I} \otimes \sigma_z$. Now suppose we prepare two electrons in the singlet state, place the two machines 1 light-year apart, and without disturbing the singlet state, move the electrons to the machines[5]. Suppose one electron is then put through its σ_z machine and returns $+1$. This tells us that the other electron will return the value -1.

One possibility is that quantum mechanics and its axioms are correct, and that the first measured electron sent the information to the second instantaneously, telling the other that it must be in the -1 state. This contradicts the principle of *locality* in special relativity, that states the maximum speed information may travel is at the (finite) speed of light, which would take 1 year to traverse a 1 light-year gap.

Alternatively, we may suppose that special-relativity is correct, and that there is some *hidden variable theory* of quantum mechanics. That is, we take the realist view that quantum mechanics is a good statistical description, but that there really was one $+1$ electron and one -1 electron the whole time, and that quantum mechanics just isn't good enough to describe it.

[5]Recall that before a quantum particle is measured, our system of axioms does not allow it to have any value. Before measuring, it is meaningless to ask what the σ_z observable of one of the electrons is. It is in a superposition of $+1$ and -1 states, and until the measurement is performed, it is decisively neither.

Theorem 4.6. *There are no local hidden variable theories consistent with quantum mechanics.*

Proof. Assume that the results of quantum mechanics are correct. Consider an arbitrary unit-vector in $\mathbb{R}^3$, $r = \sin(\theta)\cos(\phi)\hat{x} + \sin(\theta)\sin(\phi)\hat{y} + \cos(\theta)\hat{z}$, then call the spin-observable along the r direction

$$\sigma_r = \begin{pmatrix} \cos(\theta) & \sin(\theta)e^{-i\phi} \\ \sin(\theta)e^{i\phi} & -\cos(\theta) \end{pmatrix}$$

The eigenvalues of such an operator are ± 1. Let a and b be unit vectors. It can be shown by direct calculation[6] that the expectation value of the operator $\sigma_a \otimes \mathbb{I} + \mathbb{I} \otimes \sigma_b$ is $P(a, b) = -a \cdot b$.

Suppose we had devices that could measure any σ_r, and a two spin-1/2 fermion system, an electron and positron, in the singlet state. Further suppose that the system is characterized by a hidden variable, λ, (that we do not understand) that determines the outcome when the particles are measured, and also that our system is local, so particles cannot "communicate" information about their outcomes instantaneously.

Configure the machine measuring the electron spin to measure σ_a, and the machine measuring the positron to measure σ_b. Then there exist functions $A(a, \lambda)$ and $B(b, \lambda)$ which gives the actual measurement results (± 1) for the electron and positron respectively. We know in the singlet state that $A(a, \lambda) = -B(a, \lambda)$ for all λ if the detectors are aligned. If $\rho(\lambda)$ is the probability density for our hidden variable, then

$$P(a, b) = \int \rho(\lambda) A(a, \lambda) B(b, \lambda) d\lambda$$

and if the detectors are aligned we have

$$P(a, b) = -\int \rho(\lambda) A(a, \lambda) A(b, \lambda) d\lambda$$

Let c be any other unit vector. Then since $[A(b, \lambda)]^2 = 1$

$$\begin{aligned} P(a, b) - P(a, c) &= -\int \rho(\lambda)[A(a, \lambda)A(b, \lambda) - A(a, \lambda)A(c, \lambda)]d\lambda \\ &= -\int \rho(\lambda)[1 - A(b, \lambda)A(c, \lambda)]A(a, \lambda)A(b, \lambda)d\lambda \end{aligned}$$

[6] See Griffiths problem 4.50

and since $-1 \leq A(a,\lambda)A(b,\lambda) \leq 1$, and $\rho(\lambda)[1 - A(b,\lambda)A(c,\lambda)] \geq 0$, we have

$$\begin{aligned} |P(a,b) - P(a,c)| &\leq \int \rho(\lambda)[1 - A(b,\lambda)A(c,\lambda)]d\lambda \\ &= 1 + P(b,c) \end{aligned}$$

Now choose a and b to meet at 90°, choose c so that c bisects a and b at a 45° angle. Then the quantum mechanical calculations we referenced give $P(a,b) = 0$, $P(a,c) = P(b,c) = -1/\sqrt{2}$. Which contradicts the inequality for our hidden variable theory.

□

From this we conclude the following: If we really believe our universe is deterministic, then quantum mechanics must be superseded by some non-local hidden variable theory.

References:

[1] V. Bargmann. Note on wigner's theorem on symmetry operations. *Journal of Mathematical Physics*, 5(7):862–868, 1964.

[2] David S. Dummit and Richard M. Foote. *Abstract Algebra*. Wiley, 3 edition, 2004.

[3] William Fulton and Joe Harris. *Representation Theory: A First Course*. Springer-Verlag, 1991.

[4] Andrew Gleason. Measures on the closed subspaces of a hilbert space. *Indiana University Mathematics Journal*, 6(4):885–893, 1957.

[5] Roe Goodman and Nolan R. Wallach. *Symmetry, Representations, and Invariants*. Springer, 2009.

[6] David J. Griffiths. *Introduction to Quantum Mechanics*. Pearson, 2005.

[7] Brian C. Hall. *Quantum Theory for Mathematicians*. Springer, 2013.

[8] Brian C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Entroduction*. Springer, 2015.

[9] George W. Mackey. *Unitary Group Representations: in Physics, Probability, and Number Theory*. Addison-Wesley, 1989.

[10] George W. Mackey. *Mathematical Foundations of Quantum Mechanics*. Dover Publications, 2004.

[11] Michael E. Peskin and Daniel V. Schroeder. *An Introduction To Quantum Field Theory*. Westview Press Incorporated, 1995.

[12] Walter Rudin. *Real and Complex Analysis*. McGraw-Hill Series in Higher Mathematics. WCB McGraw-Hill, 1987.

[13] Terrance Tao. Noethers theorem, and the conservation laws for the euler equations. *What's new*, Mar 2014.

[14] V. S. Varadarajan. *Geometry of Quantum Theory*. Springer Science Business Media LLC, 2007.

[15] S. Weinberg. *The Quantum Theory of Fields: Volume 1: Foundations*. Cambridge University Press, 1999.

[16] Eugene P. Wigner. *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra*. Academic Press, 1959.

[17] Peter Woit. *Quantum Theory, Groups and Representations: An Introduction*.

[18] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta-decay. *The Theory of Beta-Decay*, pages 160–165, 1969.