The Pennsylvania State University

The Graduate School

Geometry of Quantum Mechanics

A Thesis in
Physics
by
Troy A. Schilling

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 1996

Abstract

While the traditional formulation of quantum mechanics is essentially linear and algebraic, that of the classical description is non-linear, and geometric; the mathematical structures involved in the respective formalisms are, it seems, radically different. The fact that such a non-linear description of nature arises, in the appropriate regime, from purely linear interactions is quite remarkable.

As we will see, however, the mathematical differences between classical and quantum mechanics are not so dramatic as they initially appear. We begin with the crucial observation that the true space of physical states of any quantum system is not a linear space, but a projective Hilbert space. The core result of this thesis is a formulation of *ordinary* quantum mechanics which is intrinsic to this projective space. We will find that the Hermitian inner-product of standard quantum mechanics induces a natural symplectic structure on the projective space; the space of quantum states is then a *phase space*. Moreover, this symplectic form plays precisely the same role as does the classical symplectic form. In particular, the time-evolution is determined by a preferred function on the (quantum) phase space. It will also be seen that quantum observables correspond to certain real-valued functions on the phase space and that the naturally defined (quantum) Poisson bracket of two such functions reflects the commutator of the associated operators. Quantum mechanics is therefore described as an infinite-dimensional Hamiltonian system; the mathematical language corresponds precisely with that of classical mechanics.

We find that the probabilistic features of quantum mechanics also admit a natural geometric description. The Hermitian inner-product also defines a Riemannian metric on the projective space, and transition amplitudes may be written in terms of the *distance* between points of the phase space. We therefore obtain a complete geometric formulation of quantum mechanics which cleanly separates the classical aspects of the formalism from those 'purely quantum mechanical' features. Such a language suggests new approaches to the study of quantization and semi-classical physics. A number of such issues are examined in detail. In particular, we lay the groundwork for the study of non-linear generalizations of quantum mechanics.

Contents

Notational Conventions and List of Symbols

Throughout the text we will use Penrose's abstract index notation. Abstract tensor indices are simply labels and contraction denotes the action of the tensor on its argument—not summation. Penrose's notation is especially powerful in the-infinite dimensional case, with which we will be concerned. In the rare event that a specific coordinate chart is chosen, tensor components are indicated with underlined indices.

Common symbols

$\pounds_X$	Lie derivative with respect to the vector field X
i_X	contraction with first index, i.e., $(i_X\omega)_a = X^b\omega_{ba}$
d	exterior derivative on differential forms
$\wedge$	exterior product on differential forms
π^*	pull-back via the mapping π
π_*	push-forward via the mapping π
$[X,Y]$	commutator of vector fields, i.e., $[X,Y] = \pounds_X Y$
$L^\sharp$	fundamental vector field generated by Lie algebra element of group acting on a manifold
$S \xrightarrow{\pi} \mathcal{P}$	a fibre bundle with bundle manifold S, base manifold $\mathcal{P}$, and projection $\pi : S \to \mathcal{P}$
$k_{(ab)}$	$= \frac{1}{2}(k_{ab} + k_{ba})$ — symmetrization
$k_{[ab]}$	$= \frac{1}{2}(k_{ab} - k_{ba})$ — skew-symmetrization

Conventions

Poisson structure	Ω^{ab} is the unique skew-symmetric covariant tensor for which $\Omega^{ac}\Omega_{bc} = \delta^a{}_b$
Poisson bracket	$\{f,g\} := (\partial_a f)\Omega^{ab}(\partial_b g) = \Omega(X_f, X_g)$
Hamiltonians	The Hamiltonian vector field generated by f is defined by $X_f{}^a := \Omega^{ab}\partial_b f$, and satisfies $i_{X_f}\Omega = \mathrm{d}f$
Kähler form	If $J^a{}_b$ is a complex structure and G_{ab} is a (real) inner-product on a Hermitian vector space, the Kähler form is defined by $\Omega_{ab} := -G_{ac}J^c{}_b$
Curvature tensor	$R_{abc}{}^d k_d = \big(\nabla_a\nabla_b - \nabla_b\nabla_a\big)k_c \quad \forall\, k_c$

List of Symbols

D_a	Levi-Cività derivative operator on $\mathcal{H}$, compatible with G
$D_a^{'}$	Levi-Cività derivative operator on $\mathcal{H}^{\times}$, compatible with $G^{'}$
∇_α	Levi-Cività derivative operator on $\mathcal{P}$, compatible with g
G_{ab}	Riemannian metric on $\mathcal{H}$
$G_{ab}^{'}$	conformally related metric on $\mathcal{H}^{\times}$, $G_{ab}^{'} = (2\hbar/r^2)G_{ab}, r^2 = G_{ab}\Psi^a\Psi^b$
$g_{\alpha\beta}$	Riemannian metric on the projective space $\mathcal{P}$
$\mathcal{H}$	Hilbert space
$\mathcal{H}^{\times}$	the 'punctured' Hilbert space, $\mathcal{H}-\{\mathbf{0}\}$
$J^a{}_b$	complex structure on $\mathcal{H}$ representing multiplication by i
$\mathcal{J}^a$	$-J^a{}_b\Psi^b$; generator of phase rotations
$\mathcal{O}_{cl}$	algebra of classical observables
$\mathcal{O}_{qu}$	algebra of quantum observables
$\mathcal{P}$	projective Hilbert space
π	projection mapping $\pi : S \to \mathcal{P}$
Π	projection mapping $\Pi : \mathcal{H}^{\times} \to \mathcal{P}$
Ψ^a	radial vector field on $\mathcal{H}$ defined by canonical identification of tangent spaces with $\mathcal{H}$ itself
S	sphere of vectors with unit norm (w.r.t. $\langle , \rangle$)
S_k	sphere of vectors with norm equal to k (w.r.t. $\langle , \rangle$)
$\mathcal{S}_p$	space of symmetry data at $p \in \mathcal{P}$
$\tilde{T}^{a_1\cdots a_m}_{b_1\cdots b_n}$	canonical lift of the tensor field $T^{\alpha_1\cdots\alpha_m}_{\beta_1\cdots\beta_n}$ on $\mathcal{P}$ to a tensor field on $\mathcal{H}^{\times}$
$\mathcal{V}$	vertical distribution of tangent spaces of $\mathcal{P}$
$\omega_{\alpha\beta}$	symplectic form on on the projective space $\mathcal{P}$
Ω_{ab}	symplectic form on $\mathcal{H}$
$\langle , \rangle$	Hermitian inner-product on $\mathcal{H}$
$\{ , \}$	Poisson bracket defined by ω
$\{ , \}_{qu}$	Poisson bracket defined by Ω
$\{ , \}_+$	Riemann bracket defined by G
$(,)$	Riemann bracket defined by g
$[,]_p$	Lie bracket on $\mathcal{S}_p$
$(,)_p$	Symmetric bracket on $\mathcal{S}_p$

Acknowledgments

Two university professors made a significant impact on my academic career. I am indebted to Robert Kadesch for sharing his joy of problem-solving and introducing me to the art of physics. Bill Sutherland encouraged me to search for *questions* as well as answers; by treating me as a sort of colleague, he provided the first glimpse of what research in mathematical physics entails and what *fun* it can be.

I am very fortunate to have been affiliated with the relativity group at Syracuse University and the Center for Gravitational Physics and Geometry at the Pennsylvania State University. Thanks to an informal, yet stimulating, environment and many talented researchers, I had opportunities to communicate with a large number of students, post-docs, local and visiting faculty members. I am particularly grateful to John Baker, Alejandro Corichi, Curt Cutler, Renate Loll, Seth Major, Don Marolf, Mark Miller, Jose Mourão, Monica Pierri, Joseph Samuel, Lee Smolin, Chopin Soo, Ranjeet Tate, Thomas Thiemann and Madhavan Varadarajan. I would also like to thank Lane Hughston for discussing many issues directly related to this work.

The members of my doctoral committee have all been of great help. Gordon Fleming, Nigel Higson, and Jorge Pullin contributed much more than a simple reading of the thesis; I am grateful to each of them for their input on a wide variety of topics. Most of all, I would like to thank Abhay Ashtekar. He is a truly inspirational and effective teacher. I am thankful for the many times he allowed me to chase my tail, for rescuing me before I had bitten it off, and for the many insights—scientific and philosophical—he has provided. He has a remarkable ability to dissect a problem into its fundamental structures without losing track of the "big picture". I hope a small fraction of that ability has rubbed off on me.

Finally, I am thankful to my close friends, Jeff McNally, April and Jon Metcalf, Subhashis Mohanty, Ric Olsen, and Brian Saltzer, for helping me to maintain my sanity and perspective over the years and to every member of my family for their unwavering encouragement. My parents, and parents in-law, have been a constant source of support; I consider myself extremely fortunate to be a member of two such loving and supportive families.

Most importantly, I would like to express the deepest gratitude to my wife, Michelle, for her patience, understanding and support. Thank you, Michelle, for allowing me to spend so much of our time on those 'important' calculations; in reality, they seem to be of little importance in comparison with our time together.

1

Introduction

Quantum mechanics was born in 1900 by an act of desperation on the part of Max Planck. The development of the subject which followed Planck's resolution of the so-called "ultraviolet catastrophe" has resulted in staggering successes. It is not only that Quantum mechanics provides a closer agreement with experiment than classical mechanics. Rather, it succeeds in regimes in which classical mechanics is not even applicable. In particular, there is no consistent classical model which accounts for the stability of the atom, let alone the detailed properties of materials. Quantum mechanics, on the other hand, provides a remarkably accurate description of the physics of molecules, atoms and nuclei; it provides a theoretical foundation for all of chemistry and, in a certain sense, the "underlying causes" of statistical mechanics. Without question, quantum mechanics has led us to the most successful scientific theories ever invented.

The arena in which quantum mechanics takes place is a Hilbert space—a *linear* space equipped with a Hermitian inner-product. Points of the Hilbert space label states of the physical system under consideration and observable quantities are represented by certain *linear* operators on this vector space. Each observable generates a flow on the Hilbert space which preserves the inner-product. The passage of time is coded by the flow of a preferred operator (the Hamiltonian); this evolution is determined by a *linear* equation. Finally, according to 'the' postulates, the notions related to measurement issues are closely tied to the *algebraic* structure. For example, measurement of a quantity represented by a particular operator will produce an element of the spectrum of that operator and the so-called "reduction of the state-vector" is described in terms of a linear projection. Quantum mechanics is, it seems, intrinsically linear and algebraic.

This picture is almost universally regarded as the fundamental description of nature, with classical mechanics arising in the appropriate regime as an approximation. The classical description is, however, intrinsically *non-linear* and *geometric*. The classical phase space, points of which label classical states, is typically a genuine manifold. This manifold is equipped with a geometric structure, called a symplectic form. In classical mechanics, observable quantities are represented by functions on the phase space and each such function determines a (Hamiltonian) vector field. Each observable then generates, by integration of the corresponding Hamiltonian vector field, a flow on the phase space which preserves the symplectic structure. As in quantum mechanics, the time

evolution is given by the flow of a preferred observable. The classical time-evolution is, however, "far from linear". In a sense, the symplectic structure is somewhat analogous to the inner-product of quantum mechanics; both are preserved by the flow induced by the observables. However, their roles seem dramatically different in other respects. While the classical symplectic structure provides a link between the observables and the dynamics, the Hermitian inner-product is the structure with which one computes transition amplitudes—entirely quantum mechanical quantities.

It is quite a mystery that this intrinsically non-linear, geometric description of nature is to arise, in the appropriate limit, from a formalism which is essentially linear and algebraic. One would expect just the reverse situation! One is more accustomed to linearized equations as approximate descriptions of some underlying non-linear phenomena. How is the geometric nature of classical mechanics encoded in the Hilbert space description of quantum mechanics?

Upon deeper reflection, ordinary quantum mechanics is not quite so linear as it is generally advertised to be. While an element of the Hilbert space determines a state of the quantum system, the converse is not true. A physical state does *not* correspond to a unique vector in the Hilbert space; multiplication of a vector by any complex number yields an element which defines the same physical state. There is an arbitrariness in the choice of the state-vector corresponding to any particular physical state; there is, in this sense, a "gauge ambiguity" inherent to the standard formalism. A physical state then corresponds to a *ray* in the Hilbert space. The true space of quantum states, which consists of these rays, is called the *projective Hilbert space*, and is a genuine manifold. It is fairly natural to ask whether there exists a formulation of quantum mechanics which is intrinsic to this true space of states—a sort of "gauge-invariant" description.

One would anticipate that such a formulation, if it exists, would rely on a geometric language. Consider, as a special case, the spin-$\frac{n}{2}$ system. The Hilbert space of this system is isomorphic to $\mathbb{C}^{n+1}$ and the projective space is $\mathbb{CP}^n$—the archetypical Kähler manifold. Since Kähler manifolds are, in particular, symplectic manifolds, certain 'classical' constructions are immediately available. For example, a function on the projective space defines a corresponding Hamiltonian vector field, there is a Poisson bracket of functions, etc. A number of very interesting questions naturally arise. What meaning can be attributed to the Hamiltonian vector fields? Recall that the classical evolution, as given by Hamilton's equation, describes the flow along the Hamiltonian vector field associated to the 'energy function'. The quantum evolution, determined by the Schrödinger equation, must define a flow on the projective space. Is this Schrödinger flow described by a Hamiltonian vector field? If such classical notions are to play a part in a geometric formulation of quantum mechanics, it seems likely that observable quantities will be represented by *functions* on the projective space. However, the space of Hermitian operators on $\mathbb{C}^{n+1}$ is finite-dimensional, while the space of smooth functions on $\mathbb{CP}^n$ is infinite-dimensional. What structure, if any, picks out a preferred class of 'observable functions'? Through answers to such questions, a geometrical formulation may provide new insights into the mathematical structures underlying even the simplest of quantum systems.

Of the current approaches to the classical limit, that which gives the strongest hint of geometric structure is in terms of coherent states [1]. Coherent states are obtained by the action (on the

vacuum, say) of the so-called dynamical group of the system [2, 3]. In the simple case of the harmonic oscillator, one may obtain a picture of the classical phase space as a subspace of the projective Hilbert space. A geometric formulation of the physics which is intrinsic to the projective space should be expected to provide a link between this algebraic procedure and the geometry of classical physics. What is the nature of the embedding of the classical phase space in the projective Hilbert space? In particular, what is the relationship between the classical symplectic structure and the geometry of the projective space? Perhaps a geometric formulation of the correspondence between classical and quantum mechanics could provide insights which are otherwise unavailable.

A deeper and more ambitious motivation for seeking a geometric description of quantum mechanics is provided by the need for a reconciliation of quantum mechanics and gravitation. The sense in which general relativity is a geometric theory goes far beyond the fact that it is a classical theory. The interaction between matter and space-time, as described by general relativity, is essentially geometric; the very solutions to Einstein's equation are, after all, space-time geometries. At the heart of gravitation is (pseudo-) Riemannian geometry. The marriage of quantum mechanics and this intrinsically geometric description of nature—the problem of *quantum gravity*—may be safely said to be deepest of all questions faced by theoretical physicists today. While the approaches to the problem of quantum gravity which have received the most attention involve the application of quantization rules to general relativity, should we not expect a modification of quantum mechanics as well? It seems fairly natural to ask whether quantum mechanics might be an approximation to an underlying, genuinely non-linear, theory, somewhat analogous to the sense in which Newtonian gravity approximates general relativity. One barrier to explorations in this direction is the fact that the linearity of the Hilbert space is *not* analogous to the linearity of Newtonian space-time; it is the fact that the Hilbert space is not the true space of physical states that presents difficulties with such a scheme. How does one "wiggle" the Hilbert space while still preserving the notion of the rays? A geometric formulation of standard quantum mechanics in a language which is not intrinsically linear seems the natural direction along which to study such 'radical' extensions of quantum mechanics.

It might be hoped that a unified description of nature would also provide a resolution of the conceptual problems introduced by quantum mechanics. The particularly enigmatic feature of the quantum mechanical formalism is the presence of two 'dynamical' processes—the U-process (unitary evolution), and the R-process (state-vector reduction), as Penrose denotes them [4]. When not subject to measurement, the quantum system evolves according to the U-process. However, when coupled to an external measuring device, the system is forced, by some mysterious mechanism, to 'collapse'; this is the R-process. One may, at times, feel that this quantum mechanical description has "no right" to work so well. Indeed, many of the founders of the formalism were, themselves, reluctant to accept the implications of their own work! Schrödinger himself is known [5] to have complained, "If we have to go on with these damned quantum jumps, then I'm sorry that I ever got involved." The phenomenon of state-vector reduction deeply troubled Einstein as well. Despite his profound respect for quantum mechanics, he firmly retained his belief in the incompleteness of the formulation. While unitary evolution is characterized by a precise mechanism, quantum

mechanics apparently lacks the ability to incorporate the means by which the reduction process takes place; the R-phenomenon is not really attributed to a 'process'. The quantum measurement problem goes deeper than this 'inadequacy' of the formalism. Any given measurement is performed by coupling a physical system to an external measuring device. However, the apparatus-plus-system itself comprises a physical system, and should also obey the laws of quantum mechanics. Since the apparatus-plus-system is not subject to external influence, it will, according to the postulates of quantum mechanics, evolve according to the U-process. Hence, a paradox. Quantum mechanics does not specify the conditions under which a measurement will even occur!

That quantum gravity—whatever that is to be—should resolve these difficulties seems a somewhat natural position. Roger Penrose has, to some degree, extended the hope of a resolution to the quantum measurement problem into a sort of 'guiding principle'. His suggestion is that gravitation provide the driving force for state-vector reduction. Unfortunately, this beautiful idea lacks a precise implementation. In order to attack the problem from this perspective, it would be useful to have a picture of quantum mechanics which is not only geometric, but also separates those mathematical structures relevant for the description of unitary evolution and state-vector reduction.

In order to briefly explore the distinction of 'U' verses 'R', consider again the example of the spin-$\frac{n}{2}$ system. The Kähler structure of $\mathbb{CP}^n$ defines not only a symplectic structure, but a Riemannian metric as well. One therefore obtains a 'quantum phase space', which is equipped with some extra structure. One can, for example, speak of the length of a curve, geodesics, etc. Since these are notions that are absent from the classical description, it is natural to ask whether these mathematical constructions are related to the physical aspects of quantum mechanics which do not have classical analogues. For example, one may wonder whether the notion of uncertainty and the probabilistic features of quantum mechanics admit a description in terms of the 'quantum metric'. Does the Riemannian structure play a role in the description of state-vector reduction? A formulation which utilizes separate geometric structures for the description of unitary evolution verses state-reduction seems likely to suggest implementations of Penrose's idea. Without a geometric description of the physics, such observations are likely to go unnoticed.

The goal of this thesis is a formulation of the postulates of standard quantum mechanics in a language which is intrinsic to the true space of states. The intent is to lay a foundation by which one may study, for example, the classical limit and non-linear modifications of quantum mechanics. The desired formalism shall be valid for the generic quantum theory; it will, in particular, encompass quantum field theory.[1] It should be emphasized that we are seeking a description of *ordinary* quantum mechanics; we introduce no new input, but merely acknowledge mathematical structures which are already inherent to the standard formalism. The difference is one of semantics, but a potentially useful one.

The description presented here allows one to adopt a viewpoint in which the Hilbert space is a *fiducial* structure, not an essential ingredient. A useful, though perhaps flattering, analogy is provided by Minkowski's geometric description of special relativity. With Minkowski's picture,

[1] While it is an important issue, we shall not consider the geometric implications of the existence of unitary representations of the Lorentz group.

one views the inertial reference frames as fiducial structures which may or may not be introduced for the description of the physics. It is not possible to over-emphasize the power of Minkowski's description; one can often replace confusing calculations involving coordinate transformations with a single diagrammatic picture. More importantly, the geometric picture of special relativity provides insights which make certain 'paradoxical' results obvious. The so-called 'twin paradox' is an example; one simply notices that a curve in Minkowski's space-time, which is nearly null, has nearly zero (proper) length. The 'paradox' is then rendered a trivial observation! The abstract description of special relativity illuminates its essential nature and provides, above all, ground-work for the formulation of general relativity. It seems quite unlikely that the general theory would have ever been brought to fruition without first obtaining Minkowski's understanding of special relativity. It is hoped that a geometric approach to quantum mechanics will provide similar insights.

The next chapter begins with a brief survey of the classical and quantum formalisms. The purpose of this chapter is two-fold; it establishes our notational conventions and provides the insights which will lead to the geometric description of quantum mechanics. Classical mechanics is discussed in terms of its natural language—that of symplectic geometry. We recall the role of the symplectic structure and the algebraic structure on the space of classical observables. Next, the essential features of quantum mechanics are reviewed and, for future reference, the postulates are stated explicitly. A notable feature of our statement of the postulates is the requirement that the observables be *bounded*. This requirement, which as we argue, is supported on physical grounds, greatly simplifies the developments to follow. Our particular emphasis is on the mathematical structures inherent to the respective formalisms. A number of heuristic similarities between the classical and quantum descriptions are then outlined. As in classical mechanics, there are *two* algebraic structures on the space of quantum observables—a Lie bracket and an Abelian product. We recall Dirac's heuristic correspondence between the Lie algebraic structures on the algebras of classical and quantum observables, and discuss the analogy between the corresponding commutative structures.

In §2.3, these heuristic observations are made concrete. We recall the fact that the quantum mechanical Hilbert space $\mathcal{H}$ is naturally described as a Kähler space; the Hermitian inner-product determines a symplectic form and a metric on the Hilbert space. We will see a sense in which it is natural to represent observables by certain functions on $\mathcal{H}$—namely, the expectation value functions corresponding to the respective operators. We find that Schrödinger's equation, when written in terms of the symplectic form, assumes the form of Hamilton's equation, the generating function being the expectation value of the Hamiltonian (in the event that the Hamiltonian is bounded). It is also seen that the commutator of linear operators on the Hilbert space is *precisely* represented by the Poisson bracket of the corresponding expectation value functions. The symplectic structure on $\mathcal{H}$ is therefore related to the dynamics, and to the algebra of observables, in the same way as is the symplectic structure of classical mechanics. The physical meaning of the metric is also explored. It is found that the notion of uncertainty, and the Jordan product, are naturally expressed in terms of the metric. At this stage, we ignore the probabilistic features of quantum mechanics.

Chapter 3 contains the core results—the formulation of quantum mechanics in a language in-

herent to the projective Hilbert space, $\mathcal{P}$. After verifying that the projective space is indeed an infinite-dimensional Hilbert manifold, we turn to its geometric structure. We see that $\mathcal{P}$ is a Kähler manifold; it then admits a symplectic structure and a Riemannian metric (both of which are strongly non-degenerate). We begin to examine the role of the symplectic structure by use of a little trickery; by viewing the space of unit vectors in $\mathcal{H}$ as a *constraint* space, we obtain a picture of the projective space as the reduced phase space of a constrained Hamiltonian system. This construction of the 'reduced phase space' suggests a natural representation of the quantum observables by certain 'observable functions' on the projective space. Using standard techniques employed for the study of constrained classical systems, we obtain an understanding of the role of the symplectic structure on $\mathcal{P}$; in particular, the quantum Poisson bracket between two observable functions precisely represents the commutator bracket between the corresponding operators on the Hilbert space.

In order to understand the meaning of the Riemannian structure on $\mathcal{P}$, as well to make closer contact with the underlying Hilbert space, it is useful to employ a particular mathematical technique for the description of Killing symmetries. This is detailed in §3.1.3. Having developed the mathematical formalism, we come to a full understanding of the role of the quantum metric in §3.2. We will have associated a real function on $\mathcal{P}$ to each quantum observable in §3.1.2, but will not have obtained a complete understanding of what characterizes these observable functions. Not all functions on $\mathcal{P}$ are associated to quantum observables; in §3.2.1 we will find that for every function on $\mathcal{P}$ whose Hamiltonian vector field also preserves the metric, there is a corresponding bounded, self-adjoint operator on $\mathcal{H}$. There is then a one-to-one correspondence between quantum observables and functions on $\mathcal{P}$ which preserve all kinematic structure.

The remainder of §3.2 is devoted to the study of those properties of quantum mechanics which do not have analogues in the classical formalism. We find, for example, that the uncertainty of an observable function is given essentially by the length, with respect to the metric, of its Hamiltonian vector field; the 'speed' of the flow generated by an observable is directly proportional to the uncertainty of that observable. An explicit expression for the Jordan product, in terms of the projective metric, is also given. Finally, we turn to the issue of measurement. Remarkably, we find that the probabilistic features of the quantum mechanical formalism are captured by the Riemannian geometry of the quantum mechanical phase space. In fact, we find that the relevant probabilities may be written in terms of the geodesic separation of points on $\mathcal{P}$. In particular, we will see a precise sense in which, as the result of measurement, the quantum system is more likely to collapse to a nearby state than a distant one.

Chapter 3 is concluded with a restatement of the postulates in geometric terms, and the example of the spin-$\frac{1}{2}$ system is presented. Quantum mechanics will have been seen to be a particularly special case of classical mechanics which incorporates an additional geometric structure. A particularly attractive feature of the formalism is the fact that the classical aspects of quantum mechanics are neatly separated, in a mathematical sense, from the purely quantum mechanical features. For example, while unitary evolution and the Poisson structure are described by the symplectic structure, the reduction process and the notion of uncertainty are described in terms of the metric.

The fact that this description of quantum mechanics possesses such obvious 'classical' features suggests a natural question: What happens if we pretend that it is a classical theory and quantize it again? This question is answered in Ch. 4. A natural guess is that the quantization procedure may be stable, i.e., that re-quantization will simply produce the quantum theory from which we started. Since the projective space admits the obvious Kähler structure, which we will have studied in detail, the obvious route to re-quantization is via geometric quantization. In §4.1. we apply this method to the 'fake classical theory' described by the symplectic structure on $\mathcal{P}$, and find that the above guess is, in fact, correct. In the next section, we observe that there are actually *many* natural symplectic structures on $\mathcal{P}$, our choice being, in a loose sense, the preferred one. We find that by quantizing all of these symplectic structures at once, one obtains the standard result of second quantization, in terms of the symmetric Fock space.

Having accomplished our main goal in Ch. 3, in Ch. 5 we make some preliminary investigations along the lines suggested in the introductory paragraphs. We begin by considering generalizations of quantum mechanics. In exploring such an issue, it is important that one understand how much of a generalization is made by altering a particular feature of the standard formalism. After all, generalizations are easy to come by; physically relevant ones are not. Therefore, in §5.1 we present a partial characterization of standard quantum mechanics in terms of the structures one is likely to modify—the manifold, the observables and the evolution. We then consider the most obvious, though perhaps not the most physical, generalization—one in which *all* functions on the projective space are allowed as observables. This is seen to correspond exactly to Weinberg's formalism [6]. The sense in which our description captures the essential nature of Weinberg's generalization is seen in §5.2.2, in which we learn that the proposal of Bialynicki-Birula [7] is, contrary to one's first impression, consistent with Weinberg's framework. The purpose of this observation is not to encourage the Bialynicki-Birula proposal, which is inconsistent with experiment [6], but to emphasize the clarity provided by the geometric approach.

In §5.3 we outline some ideas with which to study semi-classical physics; we consider only the case in which the classical phase space is linear. In this case, there is a natural fibration of the quantum phase space, which allows one to view $\mathcal{P}$ as a bundle over the classical phase space. With this picture, one can explicitly describe the kinematic classical limit; we obtain a description of the classical symplectic structure as the 'horizontal part' of the quantum symplectic structure. In fact, the horizontal spaces just alluded to are integrable, and any given leaf will be described in terms of Perelomov's generalized coherent states (as defined in [8]). In this way, we may embed the classical phase space in the quantum one in a number of ways; a particularly natural such embedding is defined by the standard coherent state space. Ideas for examining the relationship between classical and quantum dynamics are briefly explored by use of the simplest realistic example—the harmonic oscillator. Section 5.3 is concluded with a look at the Hamilton-Jacobi approach to the classical limit. We find that the 'approximate' quantum evolution determined by the mock Hamilton-Jacobi equation actually defines a generalized dynamics compatible with the Weinberg formalism.

The thesis is concluded with a brief summary in Ch. 6.

It should be stated that this work aims toward the mathematical end of the spectrum in theoretical physics. The intended audience consists of somewhat mathematically oriented theorists. Therefore, we aim for a peculiar balance; while the phenomenologist is not likely to have patience for such mathematical considerations, the mathematician may grow impatient with detailed intuitive explanations of straight-forward mathematical statements. In an attempt to achieve the right balance, every attempt is made to obtain mathematical rigor, while at the same time achieving an understanding of the physical relevance of the structures involved.

This work was intended to be an extension of a paper by Kibble [9]. However, after completing most of the work, we learned of some independent work in similar directions. The fact that the Hilbert space is naturally viewed as a symplectic space was noticed at least twenty years ago by Chernoff and Marsden [10]. In 1985, Heslot observed that quantum mechanics admits a symplectic formulation in which the phase space is the projective Hilbert space[11]; the discussion, however, was limited to the finite-dimensional case, and included neither a discussion of the role of the metric nor the probabilistic features of the formalism. Anandan and Aharonov rediscovered some of these results, and included some of the probabilistic aspects[12]. This work was also limited to the finite-dimensional regime and focused primarily on quantum evolution. Similar observations were also made by Gibbons [13] in a different context; Ch. 4 gives a complete answer to the question asked at the conclusion of Ref. [13]. A fairly complete treatment of the finite-dimensional case was given by Hughston [14]. Finally, in Ref. [15, 16] one may find discussions related to the material in Ch. 3 and §5.1.

2

Preliminaries

In order to make the above ideas more precise, we will begin with a brief review of classical mechanics in the language in which it is naturally formulated—that of symplectic geometry. We recall the roles played by the symplectic structure; its relation to the classical dynamics, and the Poisson structure on the space of classical observables determined by its inverse. In §2.2 we recall the basic features of the quantum formalism. After stating the postulates of quantum mechanics, we discuss the algebraic structures inherent to the space of quantum observables, and discuss some heuristic analogies between the classical and quantum formalisms. The main purpose of these sections is to establish our notational conventions, and provide perspective for the developments that follow. Finally, in §2.3 we add precision to the structural similarities outlined in the previous section. In particular, we find that the Schrödinger equation may be expressed in exactly the same form as Hamilton's equation, and that the commutator of quantum observables may be expressed precisely as a Poisson bracket. We also discuss the structure which does *not* have a classical analog—the Riemannian structure.

2.1 Classical Mechanics

Classical mechanics is one of the most beautiful works of science. The relationship between the Lagrangian and Hamiltonian formalisms is deep and well worth an exposition of its own. It is, in fact, difficult to find references which give justice to the structures involved in classical mechanics; Lanczos [17] provides a nice starting point. While it is tempting to provide a broad survey, we will merely review that which is essential for the development of this thesis—the Hamiltonian mechanics. For a more complete treatment, see, for example, Arnol'd [18].

The arena of Hamilton's mechanics is the classical phase space. The phase space Γ is equipped with a closed, non-degenerate 2-form, Ω_{ab}. Such a differential form is called a *symplectic form*, and (Γ, Ω) is called a *symplectic manifold*. As a consequence of the existence of Ω, Γ must be even-dimensional. Since Ω_{ab} is non-degenerate, it provides a canonical way of identifying the tangent and

cotangent spaces at any given point $p \in \Gamma$; i.e., the map

$$v^a \mapsto (i_v \Omega)_a = v^b \Omega_{ba} \tag{2.1}$$

is one-to-one and onto. It is convenient to denote the inverse mapping by the unique skew-symmetric tensor field Ω^{ab}, for which

$$\Omega_{ac}\Omega^{bc} = \delta_a{}^b. \tag{2.2}$$

With our conventions, the inverse of the mapping (2.1) may then be written

$$u_a \mapsto \Omega^{ab} u_b. \tag{2.3}$$

Given a manifold equipped with a special geometric structure, it is natural to consider diffeomorphisms which preserve that structure; in our case, such diffeomorphisms are called *canonical transformations*. Let us then suppose a vector field X generates a motion along Γ which preserves the symplectic structure, i.e.,

$$0 = \pounds_X \Omega = (i_X \mathrm{d} + \mathrm{d} i_X)\Omega = \mathrm{d} i_X \Omega. \tag{2.4}$$

The 1-form corresponding to X is then closed. Such a vector field is called a *locally Hamiltonian* vector field or, sometimes, an *infinitesimal canonical transformation*. If, in addition, $i_X \Omega$ is exact (as must be the case if, for example, Γ is simply connected), there exists a function $f : \Gamma \to \mathbb{R}$ for which

$$i_X \Omega = \mathrm{d} f, \quad \text{which implies} \quad X^a = \Omega^{ab} \partial_b f. \tag{2.5}$$

Equation (2.5) may be used to construct a vector field from any smooth function $f : \Gamma \to \mathbb{R}$. The corresponding vector field is called the *Hamiltonian vector field* associated with f, and is denoted X_f. Thus, any function f on the phase space generates a one-parameter family of canonical transformations.

It may be informative to note that the space of infinitesimal canonical transformations is therefore infinite-dimensional. This should be contrasted with the analogous situation encountered for metric manifolds, for which the space of infinitesimal symmetries (Killing vector fields) is necessarily finite-dimensional. In this sense, a symplectic structure provides less information than does a Riemannian metric. The intuitive meaning of Ω is the specification of the area of the parallelogram spanned by any two tangent vectors. A Riemannian metric determines also the lengths of the edges.

The points of Γ label states of the physical system under consideration. The time-evolution determines a flow on Γ which preserves the symplectic structure, and hence a preferred local Hamiltonian vector field X_h on Γ. We typically assume that X_h is also globally Hamiltonian (hence the notation). In this way, the dynamics is specified by a function $h : \Gamma \to R$—the classical Hamiltonian[1].

[1] For simplicity, we consider only the case for which the Hamiltonian is time-independent.

2.1.1 The algebra of classical observables

The space $\mathcal{O}_{cl}$ of classical observables consists of all smooth functions on the phase space. The concept of measurement in classical mechanics is straight-forward. If the classical system is in the physical state labeled by the point $p \in \Gamma$, then an ideal measurement of f simply produces the value $f(p)$, and does not influence the Hamiltonian evolution.

Point-wise multiplication of functions gives $\mathcal{O}_{cl}$ the structure of an associative, commutative algebra. A second important algebraic structure is provided by the Poisson bracket,

$$\begin{aligned} \{f,g\}_{cl} :&= (\partial_a f)\Omega^{ab}(\partial_b g) = X_g(f) \\ &= \Omega\left(X_f, X_g\right). \end{aligned} \tag{2.6}$$

It is easy to see that the Poisson bracket is a Lie bracket on $\mathcal{O}_{cl}$ which acts as a derivation; i.e., given $f, g, h \in \mathcal{O}_{cl}$,

$$\{f, gh\}_{cl} = g\{f,h\}_{cl} + \{f,g\}_{cl}h. \tag{2.7}$$

Since, $\pounds_{X_f}\Omega = 0$, $\pounds_X i_Y - i_Y \pounds_X = i_{[X,Y]}$, and, when acting on differential forms, the Lie derivative is given by the identity $\pounds_X = i_X \mathrm{d} + \mathrm{d}i_X$,

$$\begin{aligned} \mathrm{d}\{f,g\}_{cl} &= \mathrm{d}[\Omega(X_f, X_g)] = \mathrm{d}i_{X_g} i_{X_f}\Omega \\ &= (\pounds_{X_g} - i_{X_g}\mathrm{d}) i_{X_f}\Omega = \pounds_{X_g} i_{X_f}\Omega - i_{X_f}\pounds_{X_g}\Omega \\ &= i_{[X_g, X_f]}\Omega. \end{aligned} \tag{2.8}$$

Thus, $X_{\{f,g\}_{cl}} = -[X_f, X_g]$, so the mapping $f \mapsto X_f$ is an anti-homomorphism from the Poisson algebra of observables into the Lie algebra of smooth vector fields on Γ.

Notice that, from Eq. (2.6), if h is the classical Hamiltonian, then the rate of change of any classical observable f is given by the familiar expression,

$$\dot{f} = X_h(f) = \{f,h\}_{cl}. \tag{2.9}$$

One may recover the more familiar form of Hamilton's equations by use of the Darboux theorem [19], which guarantees the existence of local (canonical) coordinates $(q^{\underline{i}}, p_{\underline{i}})$ with respect to which the symplectic structure may be written

$$\Omega = \mathrm{d}q^{\underline{i}} \wedge \mathrm{d}p_{\underline{i}}. \tag{2.10}$$

The Poisson bracket then assumes the form

$$\{f,g\}_{cl} = \frac{\partial f}{\partial q^{\underline{i}}}\frac{\partial g}{\partial p_{\underline{i}}} - \frac{\partial f}{\partial p_{\underline{i}}}\frac{\partial g}{\partial q^{\underline{i}}}, \tag{2.11}$$

which, when combined with equation (2.9) yields the equations

$$\dot{q}^{\underline{i}} = \frac{\partial H}{\partial p_{\underline{i}}}, \qquad \dot{p}_{\underline{i}} = -\frac{\partial H}{\partial q^{\underline{i}}}.$$

The phase space described in the standard textbooks is typically a cotangent bundle $T_*\mathcal{C}$ over some configuration space $\mathcal{C}$. Any set of local coordinates $q^{\underline{i}}$ on $\mathcal{C}$ induces a natural set of coordinates $p_{\underline{i}}$ on the cotangent spaces; any covector p may be written $p = p_{\underline{i}}\mathrm{d}q^{\underline{i}}$. The symplectic structure is then defined as in Eq. (2.10). Note, however, that the abstract formulation is a slight generalization of this standard picture since there exist symplectic manifolds which are not cotangent bundles. The two-sphere is a notable example, quantization of which gives the quantum models of spin.

2.1.2 Field theory

We have tacitly restricted attention to the case in which Γ is finite-dimensional. While there are plenty of finite-dimensional classical theories with practical applications, field theory is of obvious interest as well. There are two notable complications that arise upon consideration of the infinite-dimensional case. First, a definition is in order.

Definition 2.1. A skew-symmetric tensor T_{ab} on a vector space V is called *weakly non-degenerate* if the mapping $v^a \mapsto v^b T_{ba}$ is injective. If T_{ab} is weakly non-degenerate and the above mapping is also surjective, then T is said to be *strongly non-degenerate*.

Of course, this definition may be extended to arbitrary tensors on V; in particular, one may speak of strong and weak *metrics* on V.

A strongly non-degenerate tensor then provides a canonical identification of V (a tangent space being our particular interest) and its dual (the cotangent space, in our case). Unfortunately, in classical field theory, the symplectic form is typically *weakly non-degenerate*, but not strongly so, and therefore, not invertible. While this fact introduces technical problems for the general case, our interests will allow us to avoid these problems here.

The second technicality, which we will not have the fortune of avoiding, is the fact that the Hamiltonian evolution is, in general, *not* generated by a function on the phase space. Similar to the situation in quantum mechanics, in which one considers self-adjoint operators which are defined only on dense subsets of the Hilbert space, the classical Hamiltonian function is typically only defined on a dense subset of the phase space [20]. It is the flow on Γ that is globally defined, not the generator.

2.2 Quantum Mechanics

While the classical formalism is essentially geometric in nature, the standard treatment of quantum mechanics is intrinsically algebraic. The arena in which the quantum evolution takes place is a Hilbert space, $\mathcal{H}$, and quantum observables are typically represented by self-adjoint linear operators on $\mathcal{H}$. The evolution is determined the Schrödinger equation,

$$\mathrm{i}\hbar\dot{\Psi} = \hat{H}\Psi, \tag{2.12}$$

where $\hat{H}$ is, of course, the Hamiltonian operator.

Unfortunately, there is no available structure on the space of self-adjoint operators on a Hilbert space. For example, ($-$i times) the commutator of two self-adjoint operators is *not*, in general,

self-adjoint. Clearly, life would be much simpler if we were to restrict our attention to the bounded (genuinely Hermitian) operators. Why, then, is this not the standard procedure? Presumably, the answer to this question is that the basic operators $\hat{Q}$ and $\hat{P}$, with which we are so familiar, are unbounded. If we were to eliminate the unbounded operators from the set of quantum observables, we would lose the quantum analogues of the basic classical observables.

But we would gain so much in return! Elimination of the technicalities introduced by unbounded operators not only simplifies the mathematics; it provides operational accuracy as well. To measure, say, the $\hat{Q}$ operator is to ask a question such as, "where in the world is that electron?" But the experimental physicist cannot ask that question. The experimentalist can only ask the question, "where in my *laboratory* is that electron?" The operator corresponding to this question may be written, in the Schrödinger representation, as

$$(\hat{Q}_{\rm lab}\Psi)(x) = \begin{cases} x\Psi(x), & \text{if "}x\text{ is in the lab"} \\ 0, & \text{otherwise.} \end{cases}$$

$\hat{Q}_{\rm lab}$ is clearly a bounded operator. Further, when acting on states which are concentrated in the laboratory, $\hat{Q}_{\rm lab}$ provides an excellent approximation to $\hat{Q}$. (This discussion of what is actually measured is not quite correct. The operational aspects of the type of measurements that are actually performed are briefly discussed below.) In practice, we cannot obtain arbitrarily large values as the result of an actual experiment; any given experimental setup will produce only values in a corresponding bounded interval of the real line. Thus, we will use the following:

Definition 2.2. The space $\mathcal{O}_{qu}$ of *quantum observables* consists of all bounded, self-adjoint, linear operators on the Hilbert space.

It is not immediately obvious that one may, *in practice*, measure an arbitrary observable in the laboratory. Of course, we will not pretend to resolve this issue.

It is important to note that with this definition, for most realistic systems the Hamiltonian operator will *not* be classified as an observable of the theory. This is beyond our control since the generator of the time-evolution is generally unbounded. One might complain that since energy is a physical quantity, the formalism ought to allow for its measurement. We would simply argue that any particular measurement actually corresponds to a bounded operator, which equals $\hat{H}$ in the region of Hilbert space in which the system has been prepared.

It may be of some consolation to note that this is not a special feature of quantum mechanics; it is an artifact of the infinite dimensionality of the state space. As mentioned above, for an infinite-dimensional classical theory, the Hamiltonian evolution is typically not generated by a smooth function on the phase space. In the classical (*resp.* quantum) picture, the evolution is given by a one-parameter family of symplectic (*resp.* unitary) mappings on the phase space (*resp.* Hilbert space), but the infinitesimal generator is only densely defined.

2.2.1 The postulates of quantum mechanics

For future reference, it will be convenient to have stated the postulates of quantum mechanics explicitly. Since the Hilbert space associated to essentially every physically realistic quantum theory is infinite-dimensional, a number of technical issues arise in a rigorous axiomatic approach. A treatment of these details, appropriate for the mathematically minded theorist is presented in [21]. One will find a detailed mathematical approach in [22].

Since bounded operators are defined on the entire Hilbert space, the issue of specifying domains of observables is straight-forward. However, there is an additional technical difficulty; even a bounded operator may have no eigenvalues. The $\hat{Q}_{\text{lab}}$ operator on the Hilbert space $\mathcal{H} = L^2(\mathbb{R})$ is, perhaps, the most familiar example. Any real number x_0 (corresponding to a location "in the laboratory") is "almost an eigenvalue", in the sense that

$$\text{“}(\hat{Q}_{\text{lab}}\delta)(x - x_0) = x_0\delta(x - x_0)\text{”},$$

where $\delta(x - x_0)$ is the Dirac delta 'function'. Unfortunately, since $\delta(x - x_0)$ is not an element of the Hilbert space, this equation is merely formal. Use of the projected operator $\hat{Q}_{\text{lab}}$ provides no simplification in this regard.

In infinite dimensions, one must extend the notion of the spectrum. The spectrum of an operator $\hat{A}$ is the set of complex numbers α for which the operator $\hat{A} - \alpha\mathbf{1}$ is *not* invertible (the inverse of a bounded operator, if it exists, must also be bounded). Since observables are bounded self-adjoint operators, the spectrum of any observable is a compact subset of $\mathbb{R}$ [21, Ch. 50]. Notice that an eigenvalue of $\hat{A}$ is necessarily an element of the spectrum of $\hat{A}$.

If $\mathcal{H}$ is finite-dimensional, the spectrum of an operator $\hat{A}$ consists of the eigenvalues of $\hat{A}$, and the operator may be decomposed as

$$\hat{A} = \sum_i \alpha_i \mathbb{P}_{\hat{A},\alpha_i},$$

where $\mathbb{P}_{\hat{A},\alpha_i}$ is the projection onto the eigenspace corresponding to the eigenvalue α_i. The spectral theorem for the infinite-dimensional case is quite a bit more complicated [21, Chs. 49-54]. One essentially replaces the above sum by the integral "$\hat{A} = \int \alpha \, \mathrm{d}\mathbb{P}_{\hat{A},\alpha}$", over the spectrum of $\hat{A}$. The integrand contains the 'projection' $\mathrm{d}\mathbb{P}_{\hat{A},\alpha}$, which is generally not a well-defined operator on the Hilbert space. Intuitively speaking, $\mathrm{d}\mathbb{P}_{\hat{A},\alpha}(\Psi)$ may become a Dirac delta 'function'. What *is* defined is the integral "$\mathbb{P}_{\hat{A},\Lambda} = \int_\Lambda \mathrm{d}\mathbb{P}_{\hat{A},\alpha}$" over a measurable subset, Λ, of the spectrum of $\hat{A}$. One useful peculiarity of this intuitive description is the following. Suppose Λ consists of the single point α. If α is an eigenvalue of $\hat{A}$, the above projection is simply the projection onto the corresponding eigenspace. Otherwise, it vanishes identically.

As an example, let $\mathcal{H}$ consist of the square-integrable functions on $\mathbb{R}$, and consider an operator of the form

$$(\hat{A}\Psi)(x) = f(x)\Psi(x),$$

where f is some complex-valued function. Note that for $f(x) = x$, we obtain the position operator $\hat{Q}$. The spectrum of $\hat{A}$ is the (closure of the) image of f, and the projection operator determined by the set Λ is given by

$$(\mathbb{P}_{\hat{A},\Lambda}\Psi)(x) = \begin{cases} \Psi(x), & f(x) \in \Lambda \\ 0, & \text{otherwise.} \end{cases}$$

Thus, the projection $\mathbb{P}_{\hat{A},\Lambda}$ simply kills that part of the wave function at which $f(x)$ lies outside of Λ. The complex number α is an eigenvalue of $\hat{A}$ if, for example, f assumes the constant value α on some interval of $\mathbb{R}$. For operators not of the above form, one simply goes to the representation tailored to the operator—for $\hat{P}$, this would be the momentum representation—and performs the same operation.

We are now prepared to state the postulates of quantum mechanics:

($\mathcal{H}$) *Hilbert space:* Physical states of the quantum system are in one-to-one correspondence with rays in a Hilbert space, $\mathcal{H}$.

($\mathcal{U}$) *Unitary evolution:* The system, when not subject to external influence, evolves according to the Schrödinger equation, $i\hbar\frac{\partial\Psi}{\partial t} = \hat{H}\Psi$, where $\hat{H}$ is a preferred self-adjoint operator on $\mathcal{H}$.

($\mathcal{O}$) *Observables:* Every measurable physical quantity is represented by a bounded self-adjoint operator on $\mathcal{H}$.

($\mathcal{P}$) *Probabilistic interpretation:* Let $\Lambda \subset \mathbb{R}$ be a measurable subset of the spectrum of $\hat{F}$, and suppose the system is in the state determined by the vector $\Psi \in \mathcal{H}$. The probability that measurement of $\hat{F}$ will yield an element of Λ is given by

$$P_\Psi(\Lambda) = \frac{\langle \Psi,\, \mathbb{P}_{\hat{F},\Lambda}(\Psi)\rangle}{\langle \Psi,\, \Psi\rangle}. \tag{2.13}$$

($\mathcal{R}_D$) *Reduction, discrete spectrum:* Suppose the spectrum of an observable $\hat{F} \in \mathcal{O}_{qu}$ is discrete. This spectrum provides the set of possible outcomes of the ideal measurement of $\hat{F}$. If measurement of $\hat{F}$ yields the eigenvalue λ, the state of the system immediately after the measurement is given by the associated projection, $\mathbb{P}_{\hat{F},\lambda}(\Psi)$, of the initial state Ψ.

Statement of the 'reduction postulate' ($\mathcal{R}$), for the case of the continuous spectrum, requires some further considerations. One's first thought is simply to replace the eigenvalue λ appearing in the projection with a measurable subset Λ of the spectrum. However, if measurement of $\hat{F}$ yields a value contained in Λ, then it also yields a value contained in any larger set $\Lambda' \supset \Lambda$, and the projected state $\mathbb{P}_{\hat{F},\Lambda'}(\Psi)$ is surely different than the state $\mathbb{P}_{\hat{F},\Lambda}(\Psi)$. To which state does the system collapse? This can be understood by considering what is actually done in the laboratory. In the case of a position measurement, for example, one might set up an array of detectors. A given detector corresponds to some subset Λ_i of the spectrum, which, in turn, determines the projection operator $\mathbb{P}_{\hat{Q},\Lambda_i}$. The projection operator $\mathbb{P}_{\hat{Q},\Lambda_i}$ is a Hermitian operator whose spectrum consists of the numbers 0 and

1. It is therefore an observable which asks the question, "Is the particle here, or not?" Since the Λ_i in our example are disjoint (no two detectors overlap); the operators $\mathbb{P}_{\hat{Q},\Lambda_i}$ commute amongst themselves and can be 'measured' simultaneously. One can model this situation with the operator $\hat{Q}_{\{\Lambda_i\}} := \sum_i \lambda_i \mathbb{P}_{\hat{Q},\Lambda_i}$, where λ_i is, say, the location of the center of the ith detector. The $\hat{Q}_{\{\Lambda_i\}}$ operator is a bounded, self-adjoint operator with *discrete* spectrum. In fact, the spectral theorem in infinite dimensions states that one may approximate any self-adjoint operator, as closely as possible, in just this manner. Thus, in a sense, one does not quite measure operators with continuous spectra; one measures the approximations described by the spectral theorem.

To describe the above in a more succinct way, a (bounded) self-adjoint operator $\hat{F}$ does not actually correspond to a single measurable quantity. To any measurable subset of the spectrum of $\hat{F}$ is associated a projection operator; these projection operators may be associated with measurable quantities. One may then state the 'reduction postulate' as follows:

($\mathcal{R}_C$) *Reduction, continuous spectrum:* A measurable subset Λ of the spectrum of $\hat{F}$ determines an ideal measurement that may be performed on the system. This measurement corresponds to inquiring whether the value of $\hat{F}$ lies in Λ. Immediately after this measurement, the state of the system is given by $\mathbb{P}_{\hat{F},\Lambda}(\Psi)$ or $\mathbb{P}_{\hat{F},\Lambda^c}(\Psi)$, depending on whether the result of the measurement is positive or negative, respectively.

In postulate ($\mathcal{R}_C$) above, Λ^c is the complement, in the spectrum of $\hat{F}$, of Λ. A few additional comments regarding the postulates are in order:

- If $\Psi \neq \mathbf{0}$, then for each non-zero complex number z, Ψ and $z\Psi$ correspond to the same physical state. The zero element of the Hilbert space has no physical meaning.

- Objections to the phrases "not subject to external influence", and "ideal measurement" are completely justified. We will make no attempt to justify this terminology or to pursue the philosophical implications of the postulates; we will merely examine their mathematical content.

- Although the domain of a generic self-adjoint operator, such as the Hamiltonian $\hat{H}$, is not the entire Hilbert space, one can show [10] that any self-adjoint operator generates a well-defined flow on $\mathcal{H}$.

- Let the set Λ in postulate ($\mathcal{P}$) consists of a single real number λ. If λ is an eigenvalue of $\hat{F}$, then $\mathbb{P}_{\hat{F},\Lambda}$ is the projection onto the corresponding eigenspace, and one obtains the statement found in any textbook. If λ is not an eigenvalue, the corresponding projection vanishes identically, as mentioned above.

2.2.2 The algebra of quantum observables

One of the attractive and useful features of our definition of $\mathcal{O}_{qu}$ is the fact that it possesses a nice algebraic structure. In particular, let us define

$$\{\hat{F}, \hat{G}\}_{qu} := \frac{1}{\mathrm{i}\hbar}[\hat{F}, \hat{G}], \tag{2.14}$$

where $[\hat{F}, \hat{G}]$ denotes the commutator of linear operators. It is a trivial matter to show that this defines a Lie bracket on $\mathcal{O}_{qu}$. Dirac's heuristic prescription [23, Ch. IV] of associating operators $\hat{F}$ and $\hat{G}$ to classical functions f and g, respectively, then becomes

$$\widehat{\{f, g\}}_{cl} = \{\hat{F}, \hat{G}\}_{qu}.$$

There is also a *commutative* structure—a Jordan product—on the space of observables [24], which may be defined as

$$\{\hat{F}, \hat{G}\}_{+} := \frac{1}{2}[\hat{F}, \hat{G}]_{+}, \tag{2.15}$$

where $[\hat{F}, \hat{G}]_{+} = \hat{F}\hat{G} + \hat{G}\hat{F}$ is the anti-commutator of $\hat{F}$ and $\hat{G}$. In addition to Dirac's quantization prescription, one often enforces the requirement that the Jordan product of elementary quantum operators corresponds to point-wise multiplication of the corresponding elementary classical observables [25, Ch. 10].

As in the classical description, the quantum Lie bracket acts as a derivation on the commutative algebraic structure defined by Eq. (2.15),

$$\{\hat{F}, \{\hat{G}, \hat{H}\}_{+}\}_{qu} = \{\hat{G}, \{\hat{F}, \hat{H}\}_{qu}\}_{+} + \{\{\hat{F}, \hat{G}\}_{qu}, \hat{H}\}_{+}. \tag{2.16}$$

The difference is that the commutative structure involved in classical mechanics is associative, while its quantum mechanical analogue is not. The non-associativity of the Jordan product is, however, under control; it can be written in terms of the Lie bracket as

$$\{\hat{F}, \{\hat{G}, \hat{H}\}_{+}\}_{+} - \{\{\hat{F}, \hat{G}\}_{+}, \hat{H}\}_{+} = \left(\frac{\hbar}{2}\right)^{2} \{\hat{G}, \{\hat{F}, \hat{H}\}_{qu}\}_{qu} \tag{2.17}$$

Loosely speaking, in the limit $\hbar \to 0$, the Jordan algebra "becomes associative", and a classical algebraic structure is recovered. As far as we are aware, there does not exist a clean formulation of the relationship between the classical and quantum kinematics in terms of the Jordan product. For a more precise discussion of this idea, see [26].

In this context, it is worth reflecting upon the traditional dogma which focuses on the non-commutativity which is so "fundamental" to quantum mechanics. After all, the non-commutative structure in quantum mechanics is, as Dirac pointed out long ago, strictly analogous to a non-commutative structure in classical mechanics. Moreover, the additional algebraic operation involved in quantum mechanics *is commutative*. Despite the radically different world views suggested by the respective descriptions, there are striking similarities in mathematical form. We will now begin to follow them to their logical conclusions.

2.3 The Inner-Product as a Kähler Structure

We have seen that the algebraic structures involved in quantum mechanics are quite similar to those of classical mechanics. In fact, many of these similarities can be put in a more suggestive form with a suitable description of the Hilbert space. The method is remarkably simple; describe the Hilbert space $\mathcal{H}$ using the mathematics that mirrors our intuitive description of the complex line. We typically think of $\mathbb{C}$ as a two-dimensional *real* vector space with a preferred linear operator. Multiplication by i is represented by a counter-clockwise rotation by $\pi/2$. More explicitely, the complex number $z = x + \mathrm{i}y$ is represented by the vector (x, y), so that the vector corresponding to $\mathrm{i}z$ is $(-y, x)$.

One can do exactly the same thing for any complex vector space. Thus, let us view the Hilbert space as a real vector space, on which there is a *complex structure*, J. This J is a linear operator which represents multiplication by i; therefore, $J^2 = -\mathbf{1}$. Since $\mathcal{H}$ is now viewed as a real vector space, the Hermitian inner-product is slightly unnatural. Let us then decompose it into real and imaginary parts,[2]

$$\langle \Phi, \Psi \rangle =: \frac{1}{2\hbar} G(\Phi, \Psi) + \frac{\mathrm{i}}{2\hbar} \Omega(\Phi, \Psi). \tag{2.18}$$

Since $\langle \Phi, \Psi \rangle = \overline{\langle \Psi, \Phi \rangle}$,

$$G(\Phi, \Psi) = G(\Psi, \Phi) \quad \text{and} \quad \Omega(\Phi, \Psi) = -\Omega(\Psi, \Phi). \tag{2.19}$$

Furthermore, since $\langle , \rangle$ is positive-definite, so is G;

$$G(\Psi, \Psi) \geq 0, \quad \text{and} \quad G(\Psi, \Psi) = 0 \Leftrightarrow \Psi = \mathbf{0}. \tag{2.20}$$

Thus, G is a positive-definite real inner-product on $\mathcal{H}$.

In order to display the relationship between G, Ω and J, we use the meaning of the complex structure. Since, the Hermitian inner-product respects multiplication by i, $\langle J\Phi, J\Psi \rangle = \langle \Phi, \Psi \rangle$,

$$G(J\Phi, J\Psi) = G(\Phi, \Psi) \quad \text{and} \quad \Omega(J\Phi, J\Psi) = \Omega(\Phi, \Psi), \tag{2.21}$$

i.e., J is preserved by G and Ω. Next, since $\langle \Phi, J\Psi \rangle = \mathrm{i}\langle \Phi, \Psi \rangle$, we immediately obtain the relation

$$\Omega(\Phi, \Psi) = -G(\Phi, J\Psi) = G(J\Phi, \Psi), \tag{2.22}$$

or, equivalently,

$$G(\Phi, \Psi) = \Omega(\Phi, J\Psi) = -\Omega(J\Phi, \Psi). \tag{2.23}$$

Given a complex vector space with a (real) inner-product G, one may always define a skew-symmetric bilinear form as in Eq. (2.22). This form is called the Kähler form, and the above relations define what is called a *Kähler structure*. One should note that mathematicians typically

[2] The reason for the seemingly awkward factors of $(2\hbar)^{-1}$ will become apparent below.

use a convention which differs in sign from that in Eq. (2.22). Our difference in sign may be considered a result of the physicists convention in which the Hermitian inner-product is taken to be linear in the *second* argument.

Finally, since $\mathcal{H}$ is a Hilbert space, the Hermitian inner-product is strongly non-degenerate; given any continuous linear function $A : \mathcal{H} \to \mathbb{C}$, there exists an element $\Phi_A \in \mathcal{H}$ for which $A(\Psi) = \langle \Phi_A, \Psi \rangle \; \forall \; \Psi \in \mathcal{H}$. By taking real and imaginary parts of this statement one immediately obtains the following:

Lemma 2.1. *G and Ω are strongly non-degenerate bilinear forms on $\mathcal{H}$.*

Consider, for a moment, the Kähler form Ω. By use of the canonical identification of the tangent spaces of $\mathcal{H}$ with $\mathcal{H}$ itself, we may extend Ω to a differential form on $\mathcal{H}$. This trivial operation then yields a strongly non-degenerate, closed 2-form on $\mathcal{H}$. In other words, Ω is a symplectic structure. In particular, its inverse can be used to define Poisson brackets and Hamiltonian vector fields! In the next couple sub-sections, we will see that this symplectic structure provides insight into the analogies outlined above.

2.3.1 Schrödinger's equation is *Hamilton's equation*

Let us begin with another trivial observation: A linear operator on $\mathcal{H}$ is a vector field; to the point $\Psi \in \mathcal{H}$ the operator $\hat{A}$ associates the vector $\hat{A}\Psi$. Let us consider only those operators which correspond to observables—the bounded, self-adjoint operators. The Schrödinger equation, which may be more suggestively written as

$$\dot{\Psi} = -\frac{1}{\hbar} J \hat{H} \Psi,$$

motivates us to associate, to each observable $\hat{F} \in \mathcal{O}_{qu}$, the vector field

$$Y_{\hat{F}}(\Psi) := -\frac{1}{\hbar} J \hat{F} \Psi, \tag{2.24}$$

which we will call the *Schrödinger vector field* determined by the observable $\hat{F}$. A quantum mechanical system simply evolves along the Schrödinger vector field associated with the Hamiltonian operator.

As we know from standard quantum mechanics, any self-adjoint operator generates a one-parameter family of unitary mappings on the Hilbert space. In particular, the Schrödinger vector field determined by any observable $\hat{F}$ preserves the Hermitian inner-product and, hence, both the metric G and the symplectic structure Ω. Therefore, by Eq. (2.4), $Y_{\hat{F}}$ is a locally Hamiltonian vector field, and, since $\mathcal{H}$ is a linear space, also globally Hamiltonian! In fact, the generating function has a great deal of physical content; it is simply the expectation value of $\hat{F}$.

It will be useful to see this explicitly. Given an observable $\hat{F}$, denote by $F : \mathcal{H} \to \mathbb{R}$ the expectation value function;

$$F(\Psi) := \langle \hat{F} \rangle (\Psi) = \langle \Psi, \hat{F}\Psi \rangle = \frac{1}{2\hbar} G(\Psi, \hat{F}\Psi). \tag{2.25}$$

This notation will be used throughout; the expectation value function will be denoted by "unhatting" the corresponding operator. Note that we are not dividing by the squared norm of Ψ; F is quadratic in Ψ. If η is a tangent vector at Ψ, then

$$\begin{aligned}(\mathrm{d}F)(\eta) &= \frac{\mathrm{d}}{\mathrm{d}t}\langle\Psi + t\eta,\ \hat{F}(\Psi + t\eta)\rangle\Big|_{t=0} \ = \ \langle\Psi,\ \hat{F}\eta\rangle + \langle\eta,\ \hat{F}\Psi\rangle \\ &= 2\mathrm{Re}\langle\eta,\ \hat{F}\Psi\rangle \ = \ \frac{1}{\hbar}G(\hat{F}\Psi,\ \eta) \ = \ G(JY_{\hat{F}}(\Psi),\ \eta) \\ &= \Omega(Y_{\hat{F}},\ \eta),\end{aligned} \tag{2.26}$$

where we have used, in the second line, the symmetry of $\hat{F}$, and Eqs. (2.22) and (2.24). Therefore,

$$\mathrm{d}F = i_{Y_{\hat{F}}}\Omega.$$

We then have the following:

Lemma 2.2. *The Schrödinger vector field $Y_{\hat{F}}$ determined by the observable $\hat{F} \in \mathcal{O}_{qu}$ is exactly the Hamiltonian vector field X_F generated by the expectation value of $\hat{F}$.*

This result holds without qualification in the infinite-dimensional case. As a particular consequence, in the event that $\mathcal{H}$ is finite-dimensional, the Schrödinger equation corresponds exactly to Hamilton's equation of classical mechanics, the generating function being given by the expectation value of the Hamiltonian operator.

2.3.2 The commutator is *a Poisson bracket*

The result of §2.3.1 strongly suggests a phase space formulation of quantum mechanics, in which observables are represented not by operators, but by real-valued functions, as in the classical case. We also have the obvious candidates—the expectation value functions, which have been seen to generate the motions of the corresponding operators. Let us then ask if the Lie bracket defined, in terms of the commutator, by Eq. (2.14), assumes a natural form in terms of the respective expectations. In other words, we ask whether the function

$$\left\langle \frac{1}{\mathrm{i}\hbar}[\hat{F}, \hat{K}] \right\rangle$$

is naturally expressed in terms of the functions F and K.

The calculation is remarkably simple; by use of the Hermitivity of the relevant operators and Eqs. (2.18), (2.21) and (2.24),

$$\begin{aligned}\left\langle \frac{1}{\mathrm{i}\hbar}[\hat{F},\ \hat{K}] \right\rangle(\Psi) &= \frac{1}{\mathrm{i}\hbar}\langle\Psi,\ (\hat{F}\hat{K} - \hat{K}\hat{F})\Psi\rangle \ = \ \frac{1}{\mathrm{i}\hbar}\left(\langle\hat{F}\Psi,\ \hat{K}\Psi\rangle - \langle\hat{K}\Psi,\ \hat{F}\Psi\rangle\right) \\ &= \frac{2}{\hbar}\mathrm{Im}\langle\hat{F}\Psi,\ \hat{K}\Psi\rangle \ = \ \frac{1}{\hbar^2}\Omega(\hat{F}\Psi,\ \hat{K}\Psi) \\ &= \Omega(Y_{\hat{F}},\ Y_{\hat{K}}) = \Omega(X_F,\ X_K).\end{aligned} \tag{2.27}$$

Therefore, the algebraic operation on the expectation value functions, which is induced by the commutator bracket is *exactly* a Poisson bracket! Note, however, that this is not Dirac's correspondence principle; the Poisson bracket here is the quantum one, determined by the imaginary part of the Hermitian inner-product. The Lie structure inherent to the algebra of quantum observables may be written in *precisely* the same language familiar to the classical physicist. In summary, we have seen the following:

Lemma 2.3. *Let F and K be the expectation value functions of the observables $\hat{F}$ and $\hat{K}$, respectively. Denote by H the expectation value of the Lie bracket of $\hat{F}$ and $\hat{K}$, as defined by Eq. (2.14). Then H is the Poisson bracket of F and K, with respect to the quantum symplectic structure; $H = \Omega(X_F, X_K)$.*

In what follows, we will slightly abuse notation and denote the quantum Poisson bracket by the same symbol used for the Lie bracket on $\mathcal{O}_{qu}$,

$$\{F, K\}_{qu} = \Omega(X_F, X_K). \tag{2.28}$$

This should not provide a source of confusion, as we will soon dispense with the operators and deal only with the expectation value functions.

2.3.3 Heisenberg's uncertainty principle and Riemannian geometry

The above interpretation of the imaginary part of the Hermitian inner-product has been discussed by a number of authors; see, e.g., [10, 11]. The real inner-product G, however, has received little attention (see also [14]).

The symplectic form has seen to be related to the Lie bracket on the algebra of observables. One may guess that the metric is related to the Jordan product in a somewhat similar way. That this is, in fact, the case is not difficult to see. In analogy with the question leading to Eq. (2.27), let us ask whether the function $\left\langle \frac{1}{2}[\hat{F}, \hat{K}]_+ \right\rangle$ is related, in a simple way, to the functions F and K. Just as in the previous discussion,

$$\begin{aligned} \left\langle \frac{1}{2}[\hat{F},\ \hat{K}]_+ \right\rangle (\Psi) &= \frac{1}{2}\left(\langle \hat{F}\Psi,\ \hat{K}\Psi \rangle + \langle \hat{K}\Psi,\ \hat{F}\Psi \rangle \right) = \frac{1}{2\hbar} G(\hat{F}\Psi,\ \hat{K}\Psi) \\ &= \frac{\hbar}{2} G(Y_{\hat{F}},\ Y_{\hat{K}}) = \frac{\hbar}{2} G(X_F,\ X_K). \end{aligned} \tag{2.29}$$

Therefore, up to the factor $\hbar/2$, the expectation value of the Jordan product is simply given by the (real) inner-product of the Hamiltonian vector fields associated with the expectation value functions.

As before, we will abuse notation and write

$$\{F, K\}_+ := \frac{\hbar}{2} G(X_F, X_K), \tag{2.30}$$

and will call this the *Riemann bracket* of the functions F and K. The expectation value of the Jordan product of two observables is then the Riemann bracket of the corresponding expectation value functions. This additional terminology will become useful in the next chapter.

Up to this point, we have considered those structures which have classical analogues. The measurement aspect of quantum mechanics, however, are vastly different than that of classical mechanics. In particular, the notion of uncertainty is particular to the quantum formalism. Thus, let us now turn to the consideration of the uncertainty of observables, and Heisenberg's relation. In order to avoid factors of $||\Psi||^2$, that would otherwise appear in the expressions below, we will restrict attention to the space of vectors with unit norm (with respect to the Hermitian inner-product). It is not difficult to insert appropriate scaling factors in the following equations, where they are obviously necessary.

The expression of the uncertainty of an observable $\hat{F}$, in terms of the expectation function F, is immediate:

$$(\Delta\hat{F})^2 = \langle\hat{F}^2\rangle - F^2 = \{F,F\}_+ - F^2. \tag{2.31}$$

Heisenberg's relation, as often stated, is

$$(\Delta\hat{F})^2(\Delta\hat{K})^2 \;\geq\; \left\langle -\frac{\mathrm{i}}{2}[\hat{F},\hat{K}]\right\rangle^2 \;=\; \left(\frac{\hbar}{2}\{F,K\}_{qu}\right)^2. \tag{2.32}$$

There is, however, a stronger inequality. For example, Shankar [27, Ch. 9] gives a short derivation of the following:

$$(\Delta\hat{F})^2(\Delta\hat{K})^2 \;\geq\; \left\langle -\frac{\mathrm{i}}{2}[\hat{F},\hat{K}]\right\rangle^2 + \left\langle \frac{1}{2}[\hat{F}_{\mathrm{hor}},\hat{K}_{\mathrm{hor}}]_+\right\rangle^2, \tag{2.33}$$

where $\hat{F}_{\mathrm{hor}}$ is the *nonlinear* operator defined by

$$\hat{F}_{\mathrm{hor}}(\Psi) := \hat{F}(\Psi) - F(\Psi). \tag{2.34}$$

As we will see later, $\hat{F}_{\mathrm{hor}}$ is the Ω-orthogonal part of $\hat{F}$. The inequality (2.33) then takes the form

$$(\Delta F)^2(\Delta K)^2 \;\geq\; \left(\frac{\hbar}{2}\{F,K\}_{qu}\right)^2 + \left(\{F,K\}_+ - FK\right)^2. \tag{2.35}$$

Let us briefly explore the meaning of this second term in the stronger version of the Heisenberg uncertainty relation. Given two probability distributions, ρ_1 and ρ_2, on $\mathbb{R}$, a meaningful measure of the statistical dependence of one on the other is given by the *covariance*

$$\mathrm{Cov}(\rho_1,\rho_2) := \langle\, (\rho_1 - \langle\rho_1\rangle)\,(\rho_2 - \langle\rho_2\rangle)\,\rangle\,,$$

where $\langle\rho\rangle$ denotes the mean value of ρ. Of course, a quantum observable $\hat{F}$ associates, to every (normalized) vector Ψ, a probability distribution with expectation value $\langle\hat{F}\rangle(\Psi) = F(\Psi)$. Unfortunately, the probability distributions corresponding to two different observables, $\hat{F}$ and $\hat{K}$, live on different spaces; the potential values given by measurement of $\hat{F}$ are different than those for $\hat{K}$. Therefore, the covariance, $\langle(\hat{F} - \langle\hat{F}\rangle)\cdot(\hat{K} - \langle\hat{K}\rangle)\rangle$, at Ψ does not quite make sense physically. Naively calculating this quantity would yield a complex value since $\hat{F}\hat{K}$ is, in general, not another self-adjoint

operator. However, by replacing the associative product of operators by the Jordan product, one obtains a reasonable (real) quantity, which may be viewed as the "quantum covariance",

$$\mathrm{Cov}_\Psi(\hat{F},\hat{K}) := \frac{1}{2}\left\langle [\hat{F}-F,\ \hat{K}-K]_+ \right\rangle_\Psi = \left(\{F,K\}_+ - FK\right)\big|_\Psi, \tag{2.36}$$

of $\hat{F}$ and $\hat{K}$ at Ψ. This covariance is the additional term seen in Eq. (2.35).

Notice that the usual uncertainty relation, Eq. (2.32), is saturated at Ψ only if the covariance of $\hat{F}$ and $\hat{K}$ vanishes. It is interesting that at these minimum uncertainty states, the Jordan product "acts classically", in the sense that

$$\{F,K\}_+(\Psi) = F(\Psi)K(\Psi) \quad \text{if the inequality (2.32) is saturated at } \Psi.$$

2.3.4 Remarks

Throughout this section, we have intentionally avoided use of an index notation. This has been done primarily for psychological reasons. The sight of indices, to many people, suggests the manipulation of tensor components. In infinite dimensions, summation over an infinity of components is, of course, taboo.

In this context, Penrose's abstract index notation [28, 29] is especially powerful. The superscript appearing on v^a merely indicates that v is a vector field and the contraction $T_{ab}v^aw^b$ is simply another way of writing $T(v,w)$. A caveat, however, is to be kept in mind: the tensor T_{ab} takes two vectors as its arguments. The quantity $T_{ab}G^{ab}$, which is perfectly well-defined in finite dimensions, is not defined in infinite dimensions. However $T_{ac}J^c{}_b$, for example, is well-defined; its action on the vectors v^a and w^b is simply $T_{ac}J^c{}_bv^aw^b = T(v,Jw)$. This notation can greatly simplify calculations, and will be used throughout the rest of the thesis, except in situations where the objects of interest are differential forms. In that case, we will usually use an index-free notation.

The real and imaginary parts of the Hermitian inner-product in Eq. (2.18) determine covariant tensors, G_{ab} and Ω_{ab}, which may be trivially extended to tensor fields on $\mathcal{H}$ (we will not denote these extensions with new notation). In the index notation, the first of Eqs. (2.21) becomes

$$G_{ab} = G_{cd}J^c{}_aJ^d{}_b, \tag{2.37}$$

and Eq. (2.22) becomes

$$\Omega_{ab} = -G_{ac}J^c{}_b = G_{cb}J^c{}_a. \tag{2.38}$$

Thus, Ω_{ab} is obtained by "lowering the index" of $J^a{}_b$, except for the minus sign.

A point we have overlooked is the following. An observable $\hat{F}$ determines a tensor $\hat{F}^a{}_b$ of type $(1,1)$. The requirement that $\hat{F}$ be symmetric with respect to the Hermitian inner-product corresponds to the relations

$$\begin{aligned} G_{ac}\hat{F}^c{}_b &= G_{cb}\hat{F}^c{}_a, \\ \Omega_{ac}\hat{F}^c{}_b &= \Omega_{cb}\hat{F}^c{}_a = -\Omega_{bc}\hat{F}^c{}_a. \end{aligned} \tag{2.39}$$

By lowering the first index of $\hat{F}^a{}_b$, we obtain the covariant tensor $\hat{F}_{ab}$. The first equation above merely states that this tensor is symmetric; the second states that the tensor $J^c{}_a\hat{F}_{cb} = \Omega_a{}^c\hat{F}_{cb}$ is skew-symmetric. Note that these two conditions imply that the tensor $\hat{F}^a{}_b$ respects the complex structure, i.e., $\hat{F}J = J\hat{F}$, as does any genuinely complex linear operator. Therefore, the set of observables $\mathcal{O}_{qu}$ is simply the set of symmetric tensors $\hat{F}_{ab}$ for which $J^c{}_a\hat{F}_{cb}$ is skew-symmetric.

Written in this notation, Lemma 2.2 becomes

$$Y_{\hat{F}}{}^a = X_F{}^a = \Omega^{ab}D_bF, \tag{2.40}$$

and the Poisson bracket takes the form

$$\{F,\, K\}_{qu} = (D_aF)\Omega^{ab}(D_bK), \tag{2.41}$$

where D_a is any derivative operator on $\mathcal{H}$ (for example, that compatible with G). Finally, the Riemann bracket defined by Eq. (2.30) may be expressed as

$$\{F,\, K\}_+ = \frac{\hbar}{2}G(X_F,\, X_K) = \frac{\hbar}{2}(D_aF)G^{ab}(D_bK). \tag{2.42}$$

Note that the Riemann bracket may be written with no explicit dependence on the symplectic structure.

3

Geometric Formulation of Quantum Mechanics

We now turn to the central theme of the thesis—a formulation of quantum mechanics which is intrinsic to the actual state space. Any non-zero element of the Hilbert space determines a physical state; however, the complex multiple of any vector corresponds to the same state. The space of quantum states is then the *projective Hilbert space*, points of which correspond to rays in the Hilbert space; it is a genuine manifold, which we shall denote by $\mathcal{P}$.

The first section of this chapter is devoted to the study of this projective Hilbert space and the geometric structures it possesses. Due care will be taken to insure that our discussion applies to the case of physical relevance, in which the Hilbert space is infinite-dimensional. That the projective Hilbert space is indeed a Kähler manifold is verified in §3.1. While the main goal is a description which does not refer to the underlying Hilbert space, it will be convenient, for many reasons, to have a precise way of relating structures on the projective space to corresponding structures on $\mathcal{H}$. We will present two such descriptions of $\mathcal{P}$, which will both be used to simplify the calculations that follow.

The first is quite natural from the physical point of view. While the normalization of the state vector is of no physical importance, one often calculates quantities which are *not* scale invariant (for example, the expectation value function of the previous chapter), and then restricts attention to the sphere of unit vectors. In this sense, one may view quantum mechanics as a *constrained* Hamiltonian system. The projective Hilbert space is simply the *reduced phase space* of the constrained system. This viewpoint will be especially useful in §5.2, in which we primarily discuss Weinberg's generalized quantum mechanics. Any reduced phase space—in particular the projective Hilbert space—is a symplectic manifold. The reduced phase space viewpoint will allow us to immediately interpret the symplectic structure on $\mathcal{P}$; this is done in §3.1.2. We will find that the role played by the symplectic form is precisely that of the classical formalism. In particular, the associated Poisson bracket represents the Lie structure on the space of quantum observables.

The naturality of the second description of $\mathcal{P}$ stems from a purely mathematical point of view. Since a point of $\mathcal{P}$ corresponds to a ray in the Hilbert space, it is natural to describe, say, a tensor field on the projective space as a field on $\mathcal{H}$ which is, in a sense, constant along the rays. This idea, is made precise in §3.1.3.

The Kähler structure on $\mathcal{P}$ consists of a symplectic structure and a Riemannian metric. We will have come to an understanding of the physical role played by the symplectic structure in §3.1. Section 3.2 is devoted entirely to consideration of the Riemannian structure on $\mathcal{P}$. This section begins with a discussion of the quantum observables. We will have already found that, as in classical mechanics, we may represent the observables by real-valued functions on the (quantum) phase space. Also as in classical mechanics, the observables generate flows which preserve the kinematic structures on the phase space. However, unlike the classical description, the set of such functions is a 'small' subset of the space of all smooth functions on $\mathcal{P}$.

We will find that the the Riemannian metric, which is generally missing from the classical picture, describes all features of quantum mechanics which do not have classical analogues. In particular, the notion of uncertainty and the measurement aspects of quantum mechanics are described solely by the metric. These ideas are presented in sections 3.2.2 and 3.2.3.

The results of this chapter are summarized in §3.3, in which we re-state the postulates of quantum mechanics in the geometric language. The chapter is concluded with a simple example—the spin-$\frac{1}{2}$ system.

3.1 The Quantum Phase Space

In the finite-dimensional case, the projective Hilbert space is just $\mathcal{P} = \mathbb{CP}^n$—the archetypical Kähler manifold. The notion of the Kähler manifold arises as follows. First, a *Hermitian manifold* is a complex manifold $(\mathcal{M}, \jmath^a{}_b)$, equipped with a Riemannian metric g_{ab} which respects the complex structure. This simply means that $g(v, w) = g(\jmath v, \jmath w)\ \forall v, w$. Given any such manifold, one may define a 2-form, called the Kähler form, as $\omega_{ab} := -g_{ac}\jmath^c{}_b$ (the skew-symmetry of ω is a result of the fact that g respects $\jmath$). This Kähler form will not, in general, be closed. A Kähler manifold is a Hermitian manifold whose Kähler form *is* closed. One can show that a Hermitian manifold $(\mathcal{M}, \jmath, g)$ is Kähler if and only if the complex structure is covariantly constant with respect to the Levi-Cività connection defined by g [30, Ch. 8]. As a result, the Kähler form on a Kähler manifold is not only closed, but covariantly constant.

Let us now see that $\mathcal{P}$ is, even in the infinite-dimensional regime, a Hilbert manifold—that is, a smooth manifold, modeled on a Hilbert space. Denote by $\mathcal{H}^\times$ the Hilbert space with the origin removed; $\mathcal{H}^\times := \mathcal{H} - \{\mathbf{0}\}$. Define the obvious equivalence relation,

$$\Phi \sim \Phi' \ \Leftrightarrow\ \Phi = z\Phi', \text{ for some } z \in \mathbb{C} - \{0\}, \tag{3.1}$$

so that $\mathcal{P} = \mathcal{H}^\times/\!\sim$. Any element Φ of $\mathcal{H}^\times$ determines a point $p = [\Phi]$ of the projective space, where $[\Phi]$ is the equivalence class containing Φ. Now, the set

$$V^\perp_{[\Phi]} := \{\Psi \in \mathcal{H} \mid \langle \Phi, \Psi \rangle = 0\}$$

clearly depends only on the equivalence class defined by Φ. $V^\perp_{[\Phi]}$ is a closed subspace of $\mathcal{H}$, and possesses the obvious Hermitian inner-product, induced by that on $\mathcal{H}$. Thus, to each point $p =$

$[\Phi] \in \mathcal{P}$ is associated the orthogonal subspace $V_{[\Phi]}^{\perp}$; this is the natural Hilbert space to choose for the local description of the tangent space of $\mathcal{P}$ at $p = [\Phi]$.

Given an arbitrary *normalized* element Φ of $\mathcal{H}^{\times}$, we define a mapping $\tilde{\varphi}_{\Phi} : \tilde{U}_{[\Phi]} \to V_{[\Phi]}^{\perp}$ via

$$\tilde{U}_{[\Phi]} := \mathcal{H} - V_{[\Phi]}^{\perp}, \qquad \tilde{\varphi}_{\Phi} : \Psi \mapsto \frac{1}{\langle \Phi,\ \Psi \rangle} \Psi - \Phi. \tag{3.2}$$

This map, up to re-scaling, is simply the projection onto the orthogonal subspace $V_{[\Phi]}^{\perp}$. The re-scaling has been chosen so that two elements of $[\Psi]$ are mapped to the same point. In fact, it is easy to verify the following:

Fact. *$\tilde{\varphi}_{\Phi} : \tilde{U}_{[\Phi]} \to V_{[\Phi]}^{\perp}$ is a surjective map and $\tilde{\varphi}_{\Phi}(\Psi) = \tilde{\varphi}_{\Phi}(\Psi')$ if and only if $[\Psi] = [\Psi']$.*

Proof. Let η be an arbitrary element of $V_{[\Phi]}$. Then $\eta + \Phi \in \tilde{U}_{[\Phi]}$, since $\langle \Phi,\ \eta + \Phi \rangle = 1$. Surjectivity follows from the fact that $\tilde{\varphi}_{\Phi}(\eta + \Phi) = \eta$. Next suppose $\tilde{\varphi}_{\Phi}(\Psi) = \tilde{\varphi}_{\Phi}(\Psi')$. Then $\Psi' = \frac{\langle \Phi,\ \Psi' \rangle}{\langle \Phi,\ \Psi \rangle} \Psi$, so $[\Psi'] = [\Psi]$. □

Therefore, the map $\tilde{\varphi}_{\Phi}$ induces a bijection $\varphi_{[\Phi]} : U_{[\Phi]} \to V_{[\Phi]}^{\perp}$, from the set $U_{[\Phi]} = \left\{ [\Psi] \mid \Psi \in \tilde{U}_{[\Phi]} \right\}$ onto $V_{[\Phi]}$. This bijection takes $[\Psi]$ to $\tilde{\varphi}_{[\Phi]}(\Psi)$; in particular, $\varphi_{[\Phi]}([\Phi]) = 0$. The mappings $\left(U_{[\Phi]}, \varphi_{\Phi} \right)$ will be the local charts of $\mathcal{P}$. Note that φ_{Φ} does *not* depend merely on the equivalence class of Φ.

To see that $\left\{ (U_{[\Phi]}, \varphi_{\Phi}) \mid \|\Phi\| = 1 \right\}$ forms an atlas for $\mathcal{P}$ is now fairly straight-forward. We only need to see that: *i)* $\cup\, U_{[\Phi]} = \mathcal{P}$, *ii)* φ_{Φ} is a bijection onto an open subset of $V_{[\Phi]}^{\perp}$, *iii)* $\varphi_{\Phi}\left(U_{[\Phi]} \cap U_{[\Phi']} \right)$ is open in $V_{[\Phi]}^{\perp}$, and *iv)* $\varphi_{\Phi'} \circ \varphi_{\Phi}^{-1} : \varphi_{\Phi}\left(U_{[\Phi]} \cap U_{[\Phi']} \right) \to \varphi_{\Phi'}\left(U_{[\Phi]} \cap U_{[\Phi']} \right)$ is a smooth map. The $U_{[\Phi]}$ clearly cover $\mathcal{P}$, and, by the above fact, *ii)* is obvious. The third requirement follows from the fact that each $\tilde{\varphi}_{\Phi}$ maps open sets of $\tilde{U}_{[\Phi]}$ to open sets of $V_{[\Phi]}^{\perp}$. Finally, one may write the explicit form of the transition function,

$$\varphi_{\Phi'} \circ \varphi_{\Phi}^{-1}(\eta) = \frac{\eta + \Phi}{\langle \Phi',\ \eta + \Phi \rangle} - \Phi'.$$

Since (as a result of the above fact) the denominator on the right side is non-vanishing for all $\eta \in \varphi_{\Phi}\left(U_{[\Phi]} \cap U_{[\Phi']} \right)$, this function is evidently smooth on the domain on which it is defined. The projective space is then a smooth (complex) Hilbert manifold. It may be useful for future considerations to note that if $\mathcal{H}$ is separable, then by choosing a countable orthonormal basis $\{\Phi_i\}$, one obtains a countable atlas for $\mathcal{P}$. We will not use this fact below.

The fact that the dimension of $\mathcal{P}$ is infinite poses no fatal technical problems. Some arguments are made simpler in the finite-dimensional case, by going to a local chart and working with tensor components, or by tracing tensors. These sorts of methods are not available in the infinite-dimensional regime. However, essentially all of the abstract operations available in the finite-dimensional case are available to us [31, 32]. In particular, the notions of tensor fields, the Lie derivative, and exterior differentiation of differential forms are all valid. Further, just as in finite dimensions, a strong Riemannian metric defines a torsion-free derivative operator, which may be used to parallel transport a vector along a curve. A vector at a given point of $\mathcal{P}$ determines a unique geodesic, etc. For details,

the reader should consult the references cited above. Note that all tensor fields considered below will be smooth.

One can introduce a Hermitian structure on each tangent space $T_p\mathcal{P}$ as follows. Let v and w be two tangent vectors at $p = [\Phi] \in \mathcal{P}$. These vectors may be represented by curves $p_v(t)$ and $p_w(t)$, respectively, where $p_v(0) = p = p_w(0)$. Now, let $\Psi_v(t)$ and $\Psi_w(t)$ be arbitrary lifts of $p_v(t)$ and $p_w(t)$ to the unit sphere in the Hilbert space, such that $\Psi_v(0) = \Phi = \Psi_w(0)$. That Ψ_v is a 'lift' of p_v, of course, means that $[\Psi_v(t)] = p_v(t)$; the restriction $\| \Psi_v(t) \| \equiv 1$ is merely for convenience. These lifts define two tangent vectors, V and W, at $\Phi \in \mathcal{H}$. Now define the Hermitian inner-product,

$$\langle v,\, w\rangle_p \;:=\; \langle V,\, W\rangle - \langle V,\, \Phi\rangle\langle \Phi,\, W\rangle. \tag{3.3}$$

In order to explore the consistency of Eq. (3.3), suppose $\Psi_v^{'}(t)$ and $\Psi_w^{'}(t)$ are two different lifts of the curves $p_v(t)$ and $p_w(t)$, repsectively. Then[1] $\Psi_v^{'}(t) = \mathrm{e}^{J\phi_1(t)}\Psi_v(t)$ and $\Psi_w^{'}(t) = \mathrm{e}^{J\phi_2(t)}\Psi_w(t)$, where $\phi_1(0) = \phi_2(0) := \phi$. The corresponding tangent vectors at $\Phi^{'} := \mathrm{e}^{J\phi}\Phi \in \mathcal{H}$ are then

$$V^{'} = \dot{\phi_1}(0)J\Phi^{'} + \mathrm{e}^{J\phi}V \qquad \text{and} \qquad W^{'} = \dot{\phi_2}(0)J\Phi^{'} + \mathrm{e}^{J\phi}W.$$

Inserting these into Eq. (3.3) immediately gives

$$\langle V^{'},\, W^{'}\rangle - \langle V^{'},\, \Phi^{'}\rangle\langle \Phi^{'},\, W^{'}\rangle \;=\; \langle V,\, W\rangle - \langle V,\, \Phi\rangle\langle \Phi,\, W\rangle \;=\; \langle v,\, w\rangle_p,$$

so Eq. (3.3) provides a well-defined inner-product on $T_p\mathcal{P}$. We have essentially taken the 'orthogonal part' of the Hermitian inner-product on $\mathcal{H}$.

When expressed in terms of a local chart $(U_{[\Phi]}, \varphi_\Phi)$, where $p = [\Phi]$, Eq. (3.3) assumes a somewhat illuminating form. The local representative $(\varphi_\Phi)_*(v)$ of a vector $v \in T_p\mathcal{P}$ is simply the *horizontal lift* of v, i.e. that lift of v which is orthogonal to Φ, with respect to the Hermitian inner-product on $\mathcal{H}$. The local expression of Eq. (3.3) then reduces to the inner-product on $V^{\perp}_{[\Phi]}$. In particular, $\langle , \rangle_p$ is *strongly non-degenerate*. Rather than pursuing local expressions any further, it will be more convenient to describe the geometric structures on $\mathcal{P}$ implicitly, using the ideas presented in the next two sub-sections.

3.1.1 The "constrained system" description

One of the most important concepts in modern theoretical physics is that of gauge. Loosely speaking, gauge is the ambiguity in our mathematical description of the physics under consideration, and is typically due to *constraints* on that part of the phase space accessible to the physical system. These constraints usually arise from a degeneracy in the Legendre transformation, which identifies the tangent bundle of the configuration space with its cotangent bundle. The use of the term "phase space", in this context, is a bit awkward; after all, there are points of the "phase space" which do not correspond to physical states. The terminology is, however, customary; it is probably due

[1] One can simply make the replacement $J \mapsto \mathrm{i}$ in the equations below. The notation is meant to encourage the viewpoint which describes $\mathcal{H}$ as a real vector space, equipped with a complex structure.

to the usual way of thinking of phase space as a cotangent bundle. Familiarity with the physics of constrained systems is assumed in the following discussion. For a short introduction to this important subject, refer to Dirac's little book [33]. A more modern and comprehensive treatment may be found in [34].

We have seen, in the previous chapter, that the Hilbert space is a "phase space" in that it is a symplectic space. Further, the symplectic structure interacts with the expectation value functions in the physically relevant manner. However, as in the case of a constrained classical system, there is an ambiguity in the relationship of points of this "phase space" to physical states. The traditional approach toward this ambiguity, in quantum mechanics, is to consider only state vectors which are normalized to unity; let us adopt this attitude in the most serious possible way. We therefore insist that the only physically relevant portion of the Hilbert space is that on which the *constraint function*,

$$C(\Psi) := \langle \Psi,\, \Psi \rangle - 1 \;=\; \frac{1}{2\hbar} G(\Psi,\, \Psi) - 1, \tag{3.4}$$

vanishes. The constraint space is the unit sphere,

$$S := \{ \Psi \in \mathcal{H} \mid \langle \Psi,\, \Psi \rangle \;=\; \frac{1}{2\hbar} G(\Psi,\, \Psi) \;=\; 1 \}. \tag{3.5}$$

We treat this condition, not as a convention, but as a restriction on the physically relevant part of the "phase space". As we will see, this way of thinking necessarily generalizes quantum mechanics, but ever so slightly. The state vector is now defined up to an overall phase rotation; we have *not* eliminated all of the ambiguity involved in the Hilbert space description.

Since we already know that the quantum dynamics is specified by a one-parameter family of unitary mappings on $\mathcal{H}$, the constraint is preserved by the Schrödinger evolution. Therefore, the constraint is consistent; no secondary constraints need to be introduced. In the general theory, constraints which generate motions preserving the constraint space are called *first-class* [33]. Since we have only one constraint, it is certainly first-class ($\pounds_{X_C} C = \{C,\, C\}_{qu} = 0$). In this way, we may view quantum mechanics as a constrained classical system with a single first-class constraint.

In order to explain the implications of the first-class nature of the constraint, let us take a short digression. Suppose, for simplicity, that $\mathcal{H}$ is finite-dimensional. For the moment, we forget that we are dealing with a quantum mechanical system and pretend that $(\mathcal{H}, \Omega)$ describes a classical system. Let us also imagine that the constraint of Eq. (3.4) has resulted from a Legendre transformation. In such a scenario, the Legendre transform defines the Hamiltonian function only on its image, which is the constraint space S. Strictly speaking, the Hamiltonian function is not defined off the constraint surface. Therefore, one can not differentiate it to obtain the Hamiltonian flow; the Hamiltonian evolution is ill-defined! This is the generic situation encountered with constrained Hamiltonian systems, except for our good fortune of having only a single constraint.

In order to treat the evolution, one typically extends the Hamiltonian to a neighborhood of the constraint space. Suppose H and H' are any two such extensions. Since we are only interested in generating a flow on S (which we are still pretending is a 'classical' constraint surface), it is only

necessary that these extensions are defined on a neighborhood of the constraint space. One typically employs the following terminology: If the restrictions, to S, of two functions, F and $F^{'}$, are equal,[2] we write

$$F \approx F^{'}, \tag{3.6}$$

and say that F and $F^{'}$ are *weakly equal*. Note that if $F \approx F^{'}$, then F and $F^{'}$ differ by a constraint;

$$F^{'} \approx F \quad \Rightarrow \quad F^{'} = F + \gamma C, \tag{3.7}$$

for some function γ, defined on a neighborhood of S. Finally, we may construct the Hamiltonian vector fields associated to the extended Hamiltonian functions. The important fact, implied by the seemingly innocuous Eq. (3.7), is that the flows generated by H and $H^{'}$ are related as

$$X_{H'} = X_H + \gamma X_C\,, \tag{3.8}$$

for some *arbitrary* function γ. Therefore, the sickness of the Hamiltonian dynamics is not fatal. The dynamics is not *completely* ill-defined; the arbitrary part of the evolution is determined by the constraint. In the generic situation, in which one has a number of constraints, each first-class constraint will generate such *gauge transformations*. Here, we have just a single constraint and, therefore, a single gauge freedom.

We now return from our digression, but retain the above language. Since the constraint function is first-class, its Hamiltonian vector field is to be interpreted, as explained above, as the generator of gauge transformations. This vector field is just given by

$$X_C{}^a = \Omega^{ab} D_b C = \frac{1}{\hbar}\Omega^{ab}\Psi_b = \frac{1}{\hbar}\mathcal{J}^a, \tag{3.9}$$

where $\mathcal{J}^a := -J^a{}_b\Psi^b$. Notice that $\mathcal{J}^a$ is the generator of phase rotations on $\mathcal{H}$. Therefore, the gauge transformations generated by the constraint are exactly what they ought to be; they represent the arbitrariness in our choice of phase! The projective space—the quantum phase space—may then be interpreted as the "reduced phase space" of our "constrained classical system".

Our goal of formulating quantum mechanics in a way intrinsic to the projective Hilbert space may then be described as a desire for a *reduced phase space* formulation. The projection to the reduced phase space (in the classical case, of course) is well-understood. Our situation, in which there is only one first-class constraint, is especially simple. Denote by $\pi : S \to \mathcal{P}$, the projection onto the "reduced phase space". Here, a point of $\mathcal{P}$ is interpreted as an integral curve of $\mathcal{J}$.[3] That $\mathcal{P}$ inherits the structure of a symplectic manifold is seen as follows. Begin by pulling back the symplectic structure to S, via the immersion $i : S \to \mathcal{H}$. We then obtain the two-form $i^*\Omega$ on S, which is just the restriction of Ω to vectors tangent to the constraint space. Notice that, while $i^*\Omega$ is

[2] Again, one typically assumes that F and $F^{'}$ are defined on the entire phase space; it is only necessary, however, that they be defined on an arbitrary neighborhood of the constraint surface.

[3] $\mathcal{J}$ is a vector field on $\mathcal{H}$. Since $\mathcal{J}$ is tangent to S, it trivially defines a vector field on S, which we shall simply call $\mathcal{J}$. This minor abuse of notation should not be a source of confusion.

closed ($\mathrm{d}i^*\Omega = i^*\mathrm{d}\Omega \equiv 0$), it is surely degenerate. The degeneracy of $i^*\Omega$ is, however, under control; it is degenerate only in the gauge direction, $\mathcal{J}$. Moreover, since $\mathcal{J}^a$ generates phase rotations, which preserve the Hermitian inner-product, it preserves Ω:

$$\pounds_{\mathcal{J}}(i^*\Omega) \equiv 0. \tag{3.10}$$

Since points of $\mathcal{P}$ are integral curves of $\mathcal{J}$, $\pi_*\mathcal{J} \equiv 0$. Since $\mathcal{J}$ not only preserves $i^*\Omega$, but is also the direction of degeneracy of $i^*\Omega$, there must exist a two-form ω on $\mathcal{P}$ for which $i^*\Omega = \pi^*\omega$. Further, ω is closed, since $\pi^*(\mathrm{d}\omega) = \mathrm{d}\pi^*(\omega) = \mathrm{d}i^*(\Omega) = 0$. This is, of course, the symplectic form defined by the imaginary part of $\langle\,,\rangle_p$, as in Eq. (2.18). While the physical interpretation of the Hermitian structure may not be obvious at this stage, its imaginary part may be understood immediately. This will now be explained in detail.

3.1.2 Symplectic geometry

The fact that the projective Hilbert space may be interpreted as a reduced phase space has given us a tidy way of introducing the natural symplectic structure on $\mathcal{P}$. We now describe the physical meaning of the Poisson bracket defined by ω, as well as its relationship with dynamics.

Recall that to any quantum observable $\hat{F} \in \mathcal{O}_{qu}$, we have associated the function $F(\Psi) := \langle\Psi,\, \hat{F}\Psi\rangle$ on the Hilbert space. We found that the Hamiltonian vector field generated by F is precisely the Schrödinger vector field determined by $\hat{F}$, and that the Poisson bracket corresponds exactly to the Lie bracket on $\mathcal{O}_{qu}$. These observations suggest that if there is to exist a formulation of quantum mechanics which is intrinsic to the projective Hilbert space, it will take the form of Hamilton's classical mechanics. We will now see that this this expectation is correct.

There is a very natural way of associating, to each quantum observable $\hat{F} \in \mathcal{O}_{qu}$, a real-valued function *on the projective space*. First, restrict the expectation value function F to the constraint space, thereby obtaining the function $i^*F : S \to \mathbb{R}$. Of course, this function is gauge-invariant,

$$\pounds_{\mathcal{J}}(i^*F) \equiv 0, \tag{3.11}$$

and, therefore, defines a function $f : \mathcal{P} \to \mathbb{R}$ for which $\pi^*f = i^*F$. The statement being made here is simply that, since i^*F is constant on the integral curves of $\mathcal{J}$, and $\mathcal{P}$ is the space of these integral curves, i^*F canonically defines a function f on $\mathcal{P}$. The functions on $\mathcal{P}$ that are obtained in this manner will represent the observables in the geometric formulation of quantum mechanics. Some terminology will be useful:

Definition 3.1. Let $f : \mathcal{P} \to \mathbb{R}$ be a smooth function on $\mathcal{P}$. If there exists a bounded, self-adjoint operator $\hat{F}$ on $\mathcal{H}$ for which $\pi^*f = \langle\hat{F}\rangle\Big|_S$, f is said to be an *observable function*.

Note that there is obviously a one-to-one correspondence between quantum observables and these observable functions. As will be seen in §3.2.1, the set of observable functions on $\mathcal{P}$ is a very small subset of the entire function space.

The natural question[4] which now arises is: "What is the relationship between the Hamiltonian vector fields, ${X_f}^\alpha$ (on $\mathcal{P}$) and ${X_F}^a$ (on $\mathcal{H}$)?" We would like to simply push-forward the vector field ${X_F}^a$ to $\mathcal{P}$, and compare with ${X_f}^\alpha$. In general, one can not do such a thing. Given a point $\Psi \in S$, one can push-forward (via the derivative of the map $\pi : S \to \mathcal{P}$) the vector $X_F(\Psi)$, thereby obtaining the vector $\pi_*(X_F(\Psi))$ at $\pi(\Psi)$. By doing this at each point of S, one only obtains a well-defined vector *field* on $\mathcal{P}$ if "X_F is constant on $\pi^{-1}(p)$" for each $p \in \mathcal{P}$; i.e., if $\pounds_{\mathcal{J}} X_F \equiv 0$. Fortunately, by use of Eqs. (2.8) and (3.9),

$$\pounds_{\mathcal{J}} X_F \;=\; \hbar X_{\{C,\,F\}_{qu}} \;=\; 0. \tag{3.12}$$

The last equality of Eq. (3.12) follows from the fact that X_F is tangent to the constraint surface, so that $\{C,\,F\}_{qu} = \pounds_{X_F} C = 0$. Thus, the push-forward of ${X_F}^a|_S$ to $\mathcal{P}$ is well-defined. Finally,

$$\begin{aligned} \pi^*(i_{\pi_* X_F}(\omega)) \;&=\; i_{X_F}(i^*\Omega) \;=\; i^*(i_{X_F}\Omega) \;=\; i^*(\mathrm{d}F) \\ &=\; \mathrm{d}(i^*F) \;=\; \mathrm{d}(\pi^* f) \;=\; \pi^*(\mathrm{d}f). \end{aligned} \tag{3.13}$$

Since $\pi^* v = 0 \Leftrightarrow v = 0$, $\mathrm{d}f = i_{\pi_* X_F}(\omega)$; we have just obtained the

Lemma 3.1. *Suppose $\hat{F} \in \mathcal{O}_{qu}$ and that $F : \mathcal{H} \to \mathbb{R}$ is the corresponding expectation value function. Let $f : \mathcal{P} \to \mathbb{R}$ be that function for which $\pi^* f = F|_S$. Then the flow on $\mathcal{P}$, which is induced by the Schrödinger vector field of $\hat{F}$, is precisely the Hamiltonian flow determined by f.*

Denote the Poisson bracket of two functions $f, k : \mathcal{P} \to \mathbb{R}$ by

$$\{f,\,k\} \;=\; \omega(X_f,\,X_k) \;=\; (\nabla_a f)\omega^{ab}(\nabla_b k). \tag{3.14}$$

As an immediate corollary of the above lemma, we have

Corollary 3.2. *Suppose $F, K : \mathcal{H} \to \mathbb{R}$ are expectation value functions of two quantum observables. Let $f, k : \mathcal{P} \to \mathbb{R}$ be the functions for which $\pi^* f = F|_S$ and $\pi^* k = K|_S$. Then $\pi^*\{f,\,k\} = \{F,\,K\}_{qu}|_S$.*

Proof. $\pi^*\{f,k\} = \pi^*(\omega(X_f, X_k)) = \omega(\pi_* X_F, \pi_* X_K) = \Omega(X_F, X_K)|_S$. □

Thus, to each quantum observable $\hat{F}$ is associated a real-valued function $f : \mathcal{P} \to \mathbb{R}$ on the projective Hilbert space. The mapping $\hat{F} \mapsto f$ is one-to-one and respects the Lie algebraic structure provided by the commutator and Poisson bracket on $\mathcal{P}$, respectively. The Schrödinger vector field determined by $\hat{F}$ determines a flow on $\mathcal{P}$; this flow is generated by the Hamiltonian vector field given by $f : \mathcal{P} \to \mathbb{R}$. As an example, one may consider the finite-dimensional case, in which the Schrödinger equation is then seen to determine a 'classical' Hamiltonian dynamics on the projective space.

In order to obtain these results, we have used only the analysis of constrained Hamiltonian systems; in this context, the above is very well-known. The only novelty here is the application of

[4] Tensor fields on $\mathcal{P}$ will be denoted with greek indices.

these ideas to quantum mechanics. Similar results have been obtained, in a different context, by Kibble [9], and in the finite-dimensional regime by Heslot [11], Gibbons [13] and Hughston [14].

At this point we can see the sense in which one necessarily generalizes quantum mechanics by viewing a quantum system as a constrained Hamiltonian system. Above, we started with a special set of functions on the "phase space" $\mathcal{H}$—those which are expectation values of bounded self-adjoint operators. If we had truly adopted the constrained system viewpoint, we would have noted that it is only the value of F on the constraint space that is relevant. For example, given an observable $\hat{F}$, define the function $F' := (1 - \gamma C)F$, for some arbitrarily chosen function γ on $\mathcal{H}$. Since F and F' are weakly equal, they generate the same flow on the projective space. Ordinary quantum mechanics gauge fixes the observables by specifying a preferred extension to $\mathcal{H}$ of the 'observable functions'.

3.1.3 The "Killing reduction" description

Up to this point we have considered only those aspects of the quantum mechanical formalism which have classical analogues. We have constructed the natural symplectic structure on $\mathcal{P}$ and explored its meaning, but have neglected the Riemannian inner-product G on $\mathcal{H}$. We will now investigate the meaning of the Riemannian structure.

Of course, the metric on $\mathcal{H}$ induces a Riemannian metric on the constraint space S. Let us pull back the metric to S, just as we pulled back the symplectic form. We obtain the metric i^*G. We are really interested in a formalism which is intrinsic to the projective space; therefore, we would like to somehow project this metric to $\mathcal{P}$, as we did with the symplectic structure. Unfortunately, there is an immediate problem; while i^*G is covariantly constant with respect to $\mathcal{J}$ (again, since $\mathcal{J}$ is the generator of phase rotations, which preserve the Hermitian inner-product, and hence G), i^*G is non-degenerate. There exists no covariant tensor on $\mathcal{P}$ whose pull-back, via π, is i^*G. However, there is a simple way around this problem; subtract off the $\mathcal{J} \otimes \mathcal{J}$ part of i^*G.

Define the tensor field

$$q_{ab} := (i^*G)_{ab} - \frac{1}{2\hbar}\mathcal{J}_a\mathcal{J}_b, \tag{3.15}$$

where $\mathcal{J}_a$ is obtained by "lowering the index" with the metric. Again, we view $\mathcal{J}_a$ as a tensor field on S, but do not introduce notation to distinguish it from the corresponding tensor field on $\mathcal{H}$. Since $\mathcal{J}_a\mathcal{J}^a|_S = 2\hbar$ and $\pounds_{\mathcal{J}}G = 0$, we have:

$$\pounds_{\mathcal{J}}q = 0 \qquad \text{and} \qquad \mathcal{J}^a q_{ab} = 0. \tag{3.16}$$

Thus, there exists a symmetric tensor field $g_{\alpha\beta}$ on $\mathcal{P}$ for which $\pi^*g = q$. g is a strong Riemannian metric, and is precisely ($2\hbar$ times) the real part of the Hermitian structure defined by Eq. (3.3).

Note that $S \xrightarrow{\pi} \mathcal{P}$ is a principal $U(1)$-bundle. The decomposition of i^*G into horizontal and vertical parts, given by Eq. (3.15), defines a natural connection on S; a tangent vector $v \in T_\Psi S$ is horizontal if $v^a\mathcal{J}_a = 0$. This is simply the Berry connection, in terms of which Berry's phase is naturally described [35, 36]. In Ch. 4 we will study this principal bundle from a somewhat different point of view.

We now view the projective space as the *Killing reduction* of S with respect to $\mathcal{J}$. Recall that a vector field, on a Riemannian manifold, which preserves metric is called a *Killing vector field*. As described in [37], the space of integral curves of a Killing field itself possesses the structure of a Riemannian manifold (under certain assumptions, which are satisfied in our particular case). Our metric g, as defined below Eq. (3.16), coincides with the reduced metric described by Geroch. As he shows, there is a one-to-one correspondence between tensors $T^{\alpha_1\cdots\alpha_m}{}_{\beta_1\cdots\beta_n}$ on $\mathcal{P}$ and those tensors $T^{a_1\cdots a_m}{}_{b_1\cdots b_n}$ on S for which:

$$\pounds_{\mathcal{J}} T^{a_1\cdots a_m}{}_{b_1\cdots b_n} = 0 \qquad \text{and} \tag{3.17}$$

$$\mathcal{J}_{a_i} T^{a_1\cdots a_m}{}_{b_1\cdots b_n} \;=\; 0 \;=\; \mathcal{J}^{b_j} T^{a_1\cdots a_m}{}_{b_1\cdots b_n} \quad \forall i,j. \tag{3.18}$$

The relationship between $g_{\alpha\beta}$ and q_{ab} is a particular example. Further, this correspondence respects contraction of tensor indices and exterior differentiation of differential forms (since the lift of a form is simply the pull-back via π). This makes good intuitive sense; $\mathcal{P}$ is the set of integral curves of a Killing vector field. A tangent space of $\mathcal{P}$ looks like a subspace of $T_\Psi S$ which consists of vectors orthogonal to $\mathcal{J}$. A tensor field on $\mathcal{P}$ is represented by the unique tensor field on S which is orthogonal to $\mathcal{J}$ and is constant along the integral curves of $\mathcal{J}$.

We have a one-to-one correspondence between tensor fields on $\mathcal{P}$ and a certain class of tensor fields on S. It would be convenient if, to each $T^{\alpha_1\cdots\alpha_m}{}_{\beta_1\cdots\beta_n}$ on $\mathcal{P}$, we could associate a tensor field on $\mathcal{H}$. We can almost do this; we only need to remove the origin from $\mathcal{H}$. One may proceed as follows. Let $T^{\alpha_1\cdots\alpha_m}{}_{\beta_1\cdots\beta_n}$ be an arbitrary tensor field on $\mathcal{P}$ and, as described above, let $T^{a_1\cdots a_m}{}_{b_1\cdots b_n}$ be the corresponding 'horizontal' tensor field on S. We can extend $T^{a_1\cdots a_m}{}_{b_1\cdots b_n}$ to all of $\mathcal{H}^\times = \mathcal{H} - \{\mathbf{0}\}$ by Lie dragging it along Ψ^a, where Ψ^a is the 'radial' vector field on $\mathcal{H}^\times$, defined by the canonical identification of the tangent spaces. Since $\mathcal{J}^a$ and Ψ^a commute, this extended field will continue to be preserved by $\mathcal{J}$,

Let us explain this more precisely. $(\mathcal{H}^\times, G)$ is a Riemannian manifold equipped with the Killing vector field $\mathcal{J}^a$. A second preferred vector field on $\mathcal{H}^\times$ is the radial vector field Ψ^a. The equivalence relation (3.1) defines a map $\Pi : \mathcal{H}^\times \to \mathcal{P}$. Notice that since $\mathcal{J}^a$ and Ψ^a commute, they are surface forming; a given surface defined by these two vector fields is mapped, by Π, to a single point of $\mathcal{P}$. This is just a fancy description of the two (real) dimensional rays in the Hilbert space

Geroch's description of the Killing reduced manifold extends to the case in which one has two commuting Killing vector fields. The two (commuting) Killing vector field reduction of $(\mathcal{H}^\times, G^{'}, \mathcal{J}, \Psi)$ is the manifold whose points are integral surfaces of $\mathcal{J}$ and Ψ. This is precisely our situation if we introduce the conformal metric

$$G^{'}_{ab} := \frac{2\hbar}{r^2} G_{ab}, \qquad r^2(\Psi) = \Psi^a \Psi_a. \tag{3.19}$$

That this is true is not difficult to see. Since $\mathcal{J}$ preserves both Ψ^a and G_{ab}, it obviously preserves the conformal metric $G^{'}_{ab}$. Next, since $\pounds_\Psi G_{ab} = 2G_{ab}$ and $\pounds_\Psi(r^2) = 2r^2$, Ψ^a also preserves $G^{'}_{ab}$. Therefore, any tensor field $T^{\alpha_1\cdots\alpha_m}{}_{\beta_1\cdots\beta_n}$ on $\mathcal{P}$, canonically defines a tensor field $\tilde{T}^{a_1\cdots a_m}{}_{b_1\cdots b_n}$, on

$\mathcal{H}^\times$, for which

$$\begin{aligned}
\pounds_\Psi \tilde{T}^{a_1\cdots a_m}{}_{b_1\cdots b_n} &= 0 = \pounds_\mathcal{J} \tilde{T}^{a_1\cdots a_m}{}_{b_1\cdots b_n}, \\
\Psi^{b_i} \tilde{T}^{a_1\cdots a_m}{}_{b_1\cdots b_n} &= 0 = \Psi_{a_j} \tilde{T}^{a_1\cdots a_m}{}_{b_1\cdots b_n} \qquad \forall i,j, \\
\mathcal{J}^{b_i} \tilde{T}^{a_1\cdots a_m}{}_{b_1\cdots b_n} &= 0 = \mathcal{J}_{a_j} \tilde{T}^{a_1\cdots a_m}{}_{b_1\cdots b_n} \qquad \forall i,j.
\end{aligned} \tag{3.20}$$

Conversely, any tensor field $\tilde{T}^{a_1\cdots a_m}{}_{b_1\cdots b_n}$ on $\mathcal{H}^\times$, which satisfies the requirements expressed in Eqs. (3.20) defines a tensor field on $\mathcal{P}$. This notation will be used throughout; when it is convenient, we will denote by $\tilde{T}$, the "horizontal lift" (to $\mathcal{H}^\times$) of the tensor field T (on $\mathcal{P}$). Note that the horizontal lift of a differential form is simply its pull-back via Π. Again, the mapping $T \mapsto \tilde{T}$ respects contraction of indices and exterior differentiation of differential forms.

Let us look at the obvious examples. To the observable $\hat{F}$ is associated the function F on $\mathcal{H}$, which projects to the function $f : \mathcal{P} \to \mathbb{R}$. The lift $\tilde{f}$ is, of course, the normalized expectation value function,

$$\tilde{f}(\Psi) = \frac{2\hbar}{r^2} F(\Psi) = \frac{\langle \Psi,\, \hat{F}\Psi \rangle}{\langle \Psi,\, \Psi \rangle}. \tag{3.21}$$

In particular, the restriction, to S, of $\tilde{f}$ is simply the expectation,

$$\pi^* f = \tilde{f}\Big|_S = F|_S \tag{3.22}$$

The metric g may be lifted by "subtracting off" the $\mathcal{J}\otimes\mathcal{J}$ and $\Psi\otimes\Psi$ parts of G';

$$\tilde{g}_{ab} = \frac{2\hbar}{r^2} G_{ab} - \frac{2\hbar}{r^4} \left(\Psi_a \Psi_b + \mathcal{J}_a \mathcal{J}_b \right). \tag{3.23}$$

Note that the restriction of $\tilde{g}_{ab}$ to S agrees with the q_{ab} defined above. By analogy, we may define the two-form

$$\tilde{\omega}_{ab} = \frac{2\hbar}{r^2} \Omega_{ab} - \frac{2\hbar}{r^4} \left(\mathcal{J}_a \Psi_b - \Psi_a \mathcal{J}_b \right). \tag{3.24}$$

It is a simple matter to check that $\tilde{\omega}$ satisfies the requirements in Eqs. (3.20) and that the restriction of $\tilde{\omega}$ to S agrees with $i^*\Omega$; one need only notice that $\Psi^a \mathcal{J}_a = \Psi^a \Omega_{ab} \Psi^b = 0$ and $\pounds_\Psi \Omega_{ab} = 2\Omega_{ab}$. Therefore, $\tilde{\omega}$ must be the lift of the symplectic structure ω on $\mathcal{P}$.

A tensor on $T_\Psi \mathcal{H}^\times$ is said to be *horizontal* if it is orthogonal, in each index, to the vectors $\Psi^a|_\Psi$ and $\mathcal{J}|_\Psi$. One may find it useful to note that this notion of horizontality equips $\mathcal{H}^\times$ with the structure of a principal bundle. We will not directly use this fact.

An important caveat should be noted. Suppose v^α is a vector field on $\mathcal{P}$, and that $v_\alpha = g_{\alpha\beta} v^\beta$ is its dual. We may lift both to $\mathcal{H}^\times$, obtaining the fields $\tilde{v}^a$ and $\tilde{v}_a$. Then

$$\tilde{v}_a = \tilde{g}_{ab} \tilde{v}^b = G'_{ab} \tilde{v}^b = \frac{2\hbar}{r^2} G_{ab} \tilde{v}^b; \tag{3.25}$$

the lift of the dual of v^a is obtained by lowering the index with the metric G', not G. The presence of two metrics can provide a source of confusion, as well as subtle calculational errors. We will utilize

the following conventions. The tangent and cotangent spaces of $\mathcal{P}$ will always be identified by use of the metric g, i.e., $v_\alpha = g_{\alpha\beta}v^\beta$. We will identify the tangent and cotangent spaces of $\mathcal{H}^\times$ with the metric G; given a vector field v^a, by v_a we will always mean $G_{ab}v^b$, *with the following exception.* Tensor fields, on $\mathcal{H}^\times$, which are intrinsic to $\mathcal{P}$ (i.e., lifts of tensor fields on $\mathcal{P}$) will always be denoted with a 'tilde'. In order that this notation be consistent, the one-form identified with the vector field $\tilde{v}^a$ is given by $\tilde{g}_{ab}\tilde{v}^b = G'_{ab}\tilde{v}^b$. To emphasize this seemingly minor point, note that

$$\Omega_{ab}\tilde{v}^b = \frac{2\hbar}{r^2}\Omega_a{}^b\tilde{v}_b \; ! \tag{3.26}$$

For future reference, we write the horizontal lifts of some important quantities:

$$\tilde{g}_a{}^b = G_a{}^b - \frac{1}{r^2}(\Psi_a\Psi^b + \mathcal{J}_a\mathcal{J}^b), \tag{3.27}$$

$$\tilde{\omega}_a{}^b = \Omega_a{}^b - \frac{1}{r^2}(\mathcal{J}_a\Psi^b - \Psi_a\mathcal{J}^b), \tag{3.28}$$

$$\tilde{g}^{ab} = \frac{r^2}{2\hbar}G^{ab} - \frac{1}{2\hbar}(\Psi^a\Psi^b + \mathcal{J}^a\mathcal{J}^b), \tag{3.29}$$

$$\tilde{\omega}^{ab} = \frac{r^2}{2\hbar}\Omega^{ab} - \frac{1}{2\hbar}(\mathcal{J}^a\Psi^b - \Psi^a\mathcal{J}^b). \tag{3.30}$$

Equation (3.28) gives the lift of the complex structure $\mathcal{J}^\alpha{}_\beta = \omega_\beta{}^\alpha$. Note that $\tilde{g}_a{}^b(\Psi)$ is a *horizontal projection operator* in the sense that if v^a is any vector at $\Psi \in \mathcal{H}^\times$, then $\tilde{g}_a{}^b v^a$ is orthogonal to Ψ^a and $\mathcal{J}^a$.

There is also a natural way to lift the action of the derivative operator ∇_α, associated to g. Denote by D'_a the derivative operator associated to $G' = \frac{2\hbar}{r^2}G$. Recall that g arose from the Killing reduction of the *conformal* metric G'; therefore, we expect the lift of ∇_α to be expressed in terms of D'_a. This is, in fact, the case. The horizontal lift of $\nabla_\beta T^{\alpha_1\cdots\alpha_m}{}_{\beta_1\cdots\beta_n}$ may be written as [37]

$$\tilde{\nabla}_b\tilde{T}^{a_1\cdots a_m}{}_{b_1\cdots b_n} := \tilde{g}_b{}^d\,\tilde{g}_{c_1}{}^{a_1}\cdots\tilde{g}_{c_m}{}^{a_m}\,\tilde{g}_{b_1}{}^{d_1}\cdots\tilde{g}_{b_n}{}^{d_n}D'_d\tilde{T}^{c_1\cdots c_m}{}_{d_1\cdots d_n}. \tag{3.31}$$

The covariant derivative of the tensor field $\tilde{T}$ is obtained by applying D', and then projecting all indices with the horizontal projection $\tilde{g}_a{}^b$.

Let us demonstrate the use of the derivative operator $\tilde{\nabla}_a$ with a particularly useful example. Let $f : \mathcal{P} \to \mathbb{R}$ be the projection of the expectation value function $F : \mathcal{H} \to \mathbb{R}$. First, notice that

$$\Psi_a X_F{}^a = -\mathcal{J}^b D_b f = 0 \qquad \text{and} \qquad \mathcal{J}_a X_F{}^a = \Psi^b D_b F = 2F. \tag{3.32}$$

One therefore finds,

$$\begin{aligned}\tilde{X}_f{}^a = \tilde{\omega}^{ab}D_b\tilde{f} &= \frac{r^2}{2\hbar}\Omega^{ab}D_b\frac{2\hbar F}{r^2} - \frac{1}{2\hbar}\mathcal{J}^a\Psi^b D_b\frac{2\hbar F}{r^2} \\ &= X_F{}^a - \frac{2F}{r^2}\mathcal{J}^a = X_F{}^a - \frac{\tilde{f}}{\hbar}\mathcal{J}^a.\end{aligned} \tag{3.33}$$

The dual of $\tilde{X}_f{}^a$ may be obtained by contracting with $\tilde{g}_{ab}$;

$$\tilde{X}_{f\,a} = \frac{2\hbar}{r^2}\left(X_{F\,a} - \frac{\tilde{f}}{\hbar}\mathcal{J}_a\right), \tag{3.34}$$

where, of course $X_{F\,a} = G_{ab}X_F{}^b$.

We would like to calculate (the lift of) $\nabla_\alpha X_{f\,\beta}$ by use of Eq. (3.31). In order to do so, recall that given two (torsion-free) derivative operators, such as D and D', there exists a tensor field $C^c{}_{ab}$ (symmetric in its covariant indices) for which, e.g.,

$$D'_a v_b = D_a v_b - C^c{}_{ab} v_c, \qquad D'_a v^c = D_a v^c + C^c{}_{ab} v^b. \tag{3.35}$$

The actions of D and D' on tensor fields of any type are related in a similar manner, in terms of the same tensor field, $C^c{}_{ab}$. It is a straight-forward application of the material in Appendix D of [38] to see that the tensor field which relates D and D' is given by

$$C^c{}_{ab} = \frac{1}{r^2}\left(\Psi^c G_{ab} - \Psi_a G^c{}_b - \Psi_b G^c{}_a\right). \tag{3.36}$$

Finally, using Eqs. (3.31) and (3.35), one easily finds that

$$\tilde{\nabla}_a \tilde{X}_{f\,b} = \frac{2\hbar}{r^2}\tilde{g}_a{}^c \tilde{g}_b{}^d D_c X_{F\,d} + \frac{\tilde{f}}{\hbar}\tilde{w}_{ab}. \tag{3.37}$$

Since X_F is Killing, $D_c X_{F\,d}$ is skew-symmetric. Thus, the symmetric part of $\nabla_\alpha X_{f\,\beta}$ vanishes. Thus, $\pounds_{X_f} g_{\alpha\beta} = \nabla_\alpha X_{f\,\beta} + \nabla_\beta X_{f\,\alpha} = 0$, and we have the following:

Lemma 3.3. *Let $f : \mathcal{P} \to \mathbb{R}$ be any observable function. Then the Hamiltonian vector field X_f is also a Killing vector field.*

Any observable function generates a flow on $\mathcal{P}$ defined by the Hamiltonian vector field. We have found that this flow preserves *both* of the kinematic structures inherent to the quantum phase space.

Another very useful example of Geroch's technique is a simple method by which to calculate the curvature of the reduced metric g. Computing curvature can, in general, be a daunting task. However, if one knows the curvature tensor of G', it is straight-forward to obtain the (lift of the) curvature of the reduced metric. Using a method similar to that which lead to Eq. (3.37), Geroch obtains an expression for the lift of the Riemann curvature tensor of any Killing reduced metric. Denote by $R_{\alpha\beta\gamma}{}^\delta$ the Riemann tensor on $\mathcal{P}$, which, with our conventions, is defined by the equation

$$\left(\nabla_\alpha\nabla_\beta - \nabla_\beta\nabla_\alpha\right) k_\gamma =: R_{\alpha\beta\gamma}{}^\delta k_\delta, \tag{3.38}$$

for any one-form k_α. The generalization of Geroch's expression to the case in which one divides by *two* commuting Killing vector fields is

$$\tilde{R}_{abcd} = \tilde{g}_{[a}{}^m \tilde{g}_{b]}{}^n \tilde{g}_{[c}{}^p \tilde{g}_{d]}{}^q \Big[\mathfrak{R}_{mnpq} + 2(\Psi^r\Psi_r)^{-1}(D'_m\Psi_n)(D'_p\Psi_q) + 2(\Psi^r\Psi_r)^{-1}(D'_m\Psi_p)(D'_n\Psi_q) + 2(\mathcal{J}^r\mathcal{J}_r)^{-1}(D'_m\mathcal{J}_n)(D'_p\mathcal{J}_q) + 2(\mathcal{J}^r\mathcal{J}_r)^{-1}(D'_m\mathcal{J}_p)(D'_n\mathcal{J}_q)\Big],$$

where the square brackets enclosing $[ab]$ and $[cd]$ denote skew-symmetrization, and $\mathfrak{R}$ is the curvature of G'. This is the completely general expression of the curvature R of the reduction of G' with respect

to the commuting Killing vector fields $\mathcal{J}$ and Ψ. For our particular case, $\Psi^r\Psi_r = r^2 = \mathcal{J}^r\mathcal{J}_r$, and the curvature of $G^{'}$ is, as one can easily show,

$$\mathfrak{R}_{mnpq} = \frac{2}{r^2}G_{c[a}G_{b]}{}^{d} + \frac{2}{r^4}\left[\Psi_{[a}G_{b]c}\Psi^{d} - \Psi_{[a}G_{b]}{}^{d}\Psi_{c}\right].$$

Inserting this expression into the above, one immediately obtains

$$R_{\alpha\beta\gamma}{}^{\delta} = \frac{1}{\hbar}\left(g_{\gamma[\alpha}g_{\beta]}{}^{\delta} + \omega_{\alpha\beta}\omega_{\gamma}{}^{\delta} + \omega_{\gamma[\alpha}\omega_{\beta]}{}^{\delta}\right). \tag{3.39}$$

A Kähler manifold whose curvature tensor assumes this special form is said to be of *constant holomorphic sectional curvature* [39]. Such manifolds are complex analogues of manifolds of constant scalar curvature. Equation (3.39) will be used extensively in §5.1.

Before discussing the role of the metric in quantum mechanics, let us make one further mathematical note. We have found that the projective space is an infinite-dimensional complex manifold. We have also found that this manifold is equipped with two geometric structures—a Riemannian metric and a symplectic form, both of which are strongly non-degenerate. It is a simple matter to see that $\mathcal{P}$ is, indeed, a Kähler manifold. For example, one may apply the above technique to find that the complex structure, $\jmath^{\alpha}{}_{\beta} = -\omega^{\alpha}{}_{\beta} = \omega_{\beta}{}^{\alpha}$, is killed by the derivative operator ∇_{α} (and, hence, so is ω_{ab}). This could have been seen just prior to §3.1.1, by working in a local chart. Our circuitous route has, however, produced an understanding of the meaning of the symplectic structure, as well as some analytic tools to be used below.

3.2 Riemannian Geometry

We now explore the role played by the Riemannian metric g. Our first task is the characterization of the observable functions. In §3.2.1, we will find that the converse of Lemma 3.3 also holds; a function is observable if and only if its Hamiltonian vector field preserves the metric. Next, in §3.2.2, we explore the algebraic structure on the space of observables in greater detail. Naturally, from Eq. (2.29), we expect the Jordan product on the space of quantum observables to be expressed in terms of the metric. We will find, however, that it is the quantum covariance written in Eq. (2.36) that corresponds to the Riemann bracket on $\mathcal{P}$. In §3.2.3, we discuss the geometric description of the probabilistic features of quantum mechanics. As we will see, these purely quantum aspects of the formalism are described, in a surprisingly natural way, in terms of the Riemannian structure on the projective Hilbert space.

3.2.1 Quantum observables

To each quantum observable has been associated a smooth function on the projective Hilbert space. These functions, for obvious reasons, have been called observable functions. We have seen that the the algebraic structures on $\mathcal{O}_{qu}$ induce fairly natural operations on the space of observable functions; however, we don't quite know, at this point, what characterizes this set of observable functions. While Definition 3.1 is straight-forward enough, it would be nice to have a characterization of the

observable functions which is intrinsic to the quantum phase space. How does one tell whether a given (smooth) function $f : \mathcal{P} \to \mathbb{R}$ is observable?

We have seen, in Eq. (3.37), that the Hamiltonian vector field of any observable function is Killing. Let us recall an important property of Killing vector fields. Let X^α be any Killing vector field on $\mathcal{P}$. Then $\nabla_\alpha X_\beta + \nabla_\beta X_\alpha = 0$, so $K_{\alpha\beta} := \nabla_\alpha X_\beta$ defines a two-form on $\mathcal{P}$. It is not difficult to show that $\nabla_\alpha K_{\beta\gamma} = R_{\gamma\beta\alpha}{}^\delta X_\delta$ (see [38]). The implication of this fact is the following. Let p_0 and p_1 be two arbitrary points of $\mathcal{P}$ and suppose $p(t)$ is any smooth curve for which $p_0 = p(0)$ and $p_1 = p(1)$. Then,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} X_\alpha &= V^\gamma \nabla_\gamma X_\alpha = V^\gamma K_{\gamma\alpha} \quad \text{and} \\ \frac{\mathrm{d}}{\mathrm{d}t} K_{\alpha\beta} &= V^\gamma \nabla_\gamma K_{\alpha\beta} = R_{\beta\alpha\gamma}{}^\delta X_\delta, \end{aligned} \tag{3.40}$$

where $V^a(t)$ is the tangent vector to the curve $p(t)$. Therefore, given only $X^\alpha|_{p_0}$ and $K_{\alpha\beta}|_{p_0}$, one may integrate the above system of ordinary differential equations along any curve connecting p_0 and p_1 to obtain $X^\alpha|_{p_1}$ and $K_{\alpha\beta}|_{p_1}$. Thus, Killing vector fields are extremely rigid in that such a vector field is completely determined by its value and first derivative at a single point. In fact, any finite-dimensional Riemannian manifold admits only a finite number of linearly independent Killing vector fields. Note that arbitrary data $(X^\alpha, K_{\alpha\beta})$ at p_0 is, in general, not integrable, in the sense that the tensors at p_1 obtained by the above procedure will *not* be independent of the chosen curve.

Now suppose the Killing vector field above is that generated by the observable function f; below, it will be understood that $X = X_f$. Note that, in this case, $\omega_\alpha{}^\gamma K_{\gamma\beta} = -\omega_\alpha{}^\gamma \nabla_\beta X_\gamma = \nabla_\alpha \nabla_\beta f$, so $K_{\alpha\beta}$ satisfies the additional property that $\omega_\alpha{}^\gamma K_{\gamma\beta}$ is symmetric (it then defines a bounded, skew-self-adjoint operator on the tangent space $T_{p_0}\mathcal{P}$). We may then supplement Eqs. (3.40) with the defining equation

$$\nabla_\alpha f = \omega_{\gamma\alpha} X^\gamma. \tag{3.41}$$

Using the same reasoning as above, we see that *any observable function is completely determined by its value and first two derivatives at a single point.* In this sense, the set of observable functions is an extremely "small" subset of the set of smooth functions on $\mathcal{P}$. This should not be terribly surprising since in the finite-dimensional case, the function space is infinite-dimensional, while the dimension of the space of Hermitian operators is finite. The following definition will be useful:

Definition 3.2. For each point $p \in \mathcal{P}$, Let $\mathcal{S}_p$ consist of all triples, $(\lambda, X_\alpha, K_{\alpha\beta})$, where λ is a real number, X_α is a covector at p, and $K_{\alpha\beta}$ is a two-form at p for which $\omega_\alpha{}^\gamma K_{\gamma\beta} = \omega_\beta{}^\gamma K_{\gamma\alpha}$. We call $\mathcal{S}_p$ the *algebra of symmetry data at* p.

The reason for our use of the term *algebra* in the above definition will become clear shortly.

Let us obtain an understanding of the mathematical meaning of the symmetry data defined by an observable function. Suppose $\hat{F}$ is a quantum observable, and that f is the corresponding observable function on $\mathcal{P}$. Again, denote by X^α the Hamiltonian vector field of f and put $K_{\alpha\beta} = \nabla_\alpha X_\beta$. Of course, the values of (f, X, K) at a point $p \in \mathcal{P}$ define an element of $\mathcal{S}_p$. Now, recall from the

conclusion of Ch. 2 that $\Omega_{ca}\hat{F}^c{}_b$ is a skew-symmetric tensor on $\mathcal{H}$, and therefore trivially defines a two-form on $\mathcal{H}$. In fact, as is very easy to see,

$$\frac{1}{\hbar}\Omega_{ca}\hat{F}^c{}_b = D_a X_{F\,b} = \frac{1}{2}\left(\mathrm{d}X_F\right)_{ab}\,; \tag{3.42}$$

$\frac{1}{\hbar}\Omega_{ca}\hat{F}^c{}_b$ is to $X_F{}^a$ as $K_{\alpha\beta}$ is to X^α. On the other hand, from Eq. (3.34), $X_{F\,a} = \frac{r^2}{2\hbar}\tilde{X}_a + \frac{\tilde{f}}{\hbar}\mathcal{J}_a$. Applying the exterior derivative to this equation and comparing with Eq. (3.42) gives, after 'simplification',

$$\frac{1}{\hbar}\Omega_{ca}\hat{F}^c{}_b = \frac{r^2}{2\hbar}\left(\tilde{K} - \frac{\tilde{f}}{\hbar}\tilde{\omega}\right) + \frac{\tilde{f}}{\hbar r^2}\Psi\wedge\mathcal{J} + \frac{1}{2\hbar}\Psi\wedge\tilde{X} + \frac{1}{2\hbar}\left(i_{\tilde{X}}\tilde{\omega}\right)\wedge\mathcal{J}, \tag{3.43}$$

where $\tilde{X}$, Ψ and $\mathcal{J}$ are viewed as one-forms on $\mathcal{H}^\times$. Note that, while $\mathrm{d}X_F$ is well-defined on all of $\mathcal{H}$, the individual terms in Eq. (3.43) are badly behaved at the origin.

In the finite-dimensional case, one may verify that the number of independent Hermitian operators on $\mathcal{H}$ is equal to the dimension of $\mathcal{S}_p$; hence the algebra of observables is isomorphic to the algebra of symmetry data at any point. The extension of this result to the infinite-dimensional case is slightly more subtle and is provided by

Theorem 3.4. *For any element (λ, X, K) of $\mathcal{S}_p$, there exists a smooth a function $f : \mathcal{P}\to\mathbb{R}$ such that $f(p) = \lambda$, $X_{f\,\alpha}\big|_p = X_\alpha$ and $\nabla_\alpha X_{f\,\beta}\Big|_p = K_{\alpha\beta}$.*

Proof. Let Φ be an arbitrary point in $\Pi^{-1}(p)$, and let $\tilde{f}$, $\tilde{X}_a$ and $\tilde{K}_{ab}$ be the horizontal lifts of λ, X_α and $K_{\alpha\beta}$, respectively, to the point Φ. Define the tensor

$$L := \frac{r^2(\Phi)}{2\hbar}\left(\tilde{K} - \frac{\tilde{f}}{\hbar}\,\tilde{\omega}|_\Phi\right) + \frac{\tilde{f}}{\hbar r^2(\Phi)}\,(\Psi\wedge\mathcal{J})|_\Phi + \frac{1}{2\hbar}\Psi|_\Phi\wedge\tilde{X} + \frac{1}{2\hbar}\,\left(i_{\tilde{X}}\tilde{\omega}|_\Phi\right)\wedge\mathcal{J}|_\Phi$$

at the point Φ. By the canonical identification of tangent spaces of $\mathcal{H}$, we may view L as a constant tensor *field* on $\mathcal{H}$ or, equivalently, as a tensor on the *vector space* $\mathcal{H}$. By definition, L is skew-symmetric. Secondly, it is a trivial matter to show that $\Omega_a{}^c L_{cb} = \Omega_b{}^c L_{ca}$. Therefore, by the comments following Eq. (2.39), $\hat{F}^a{}_b := \hbar\Omega^{ac}L_{cb}$ defines a bounded, self-adjoint operator on $\mathcal{H}$ (compare with Eq. (3.43)). This self-adjoint operator defines an observable function $f : \mathcal{P}\to\mathbb{R}$. The value of f and its first two derivatives determine an element of $\mathcal{S}_p$; we simply check that this is precisely the original (λ, X, K), which was chosen at the beginning.

One then only needs to show the following (c.f. Eqs. (3.21), (3.33) and (3.37)): *i)* $\frac{2\hbar}{r^2(\Phi)}F(\Phi) = \tilde{f}$, *ii)* $X_F{}^a|_\Phi - \frac{\tilde{f}}{\hbar}\,\mathcal{J}^a|_\Phi = \tilde{X}^a$ and *iii)* $\frac{2\hbar}{r^2(\Phi)}\tilde{g}_a{}^c\tilde{g}_b{}^d\,D_c X_{F\,d}|_\Phi + \frac{\tilde{f}}{\hbar}\,\tilde{\omega}_{ab}|_\Phi = \tilde{K}_{ab}$, where F is, of course, the expectation value of $\hat{F}$. Each of these follows with trivial calculation. □

We have seen, in Lemma 3.3, that the Hamiltonian vector field of any observable function is Killing. Conversely, if the Hamiltonian vector field of $f : \mathcal{P}\to\mathbb{R}$ is also a Killing field, then its first two derivatives at any point p determine an element of $\mathcal{S}_p$. Therefore, by the above proof, f must be the expectation value of some bounded, self-adjoint operator on $\mathcal{H}$; this is summarized by

Corollary 3.5. *A smooth function $f : \mathcal{P} \to \mathbb{R}$ is an observable function if and only if its Hamiltonian vector field is Killing.*

Establishing this corollary was the sole purpose of this sub-section. The algebra of quantum observables $\mathcal{O}_{qu}$ is isomorphic to the set of smooth functions on $\mathcal{P}$ whose Hamiltonian vector fields are infinitesimal symmetries of the available kinematic structure, which is in turn isomorphic to $\mathcal{S}_p$, for any $p \in \mathcal{P}$. Incidentally, it is a simple matter to derive expressions for the algebraic operations on $\mathcal{S}_p$ which are induced by the Lie bracket and Jordan product on $\mathcal{O}_{qu}$; see §5.1.

3.2.2 Uncertainty revisited

We begin by expressing the Jordan product in a way which is intrinsic to the projective space. We found, in §2.3, that the Jordan product of $\hat{F}$ and $\hat{K}$ corresponds to the Riemann bracket of their expectation value functions. Of course, the Riemannian metric on $\mathcal{P}$ also defines the Riemann bracket of two functions $f, k : \mathcal{P} \to \mathbb{R}$, which we shall denote as

$$(f, k) := \frac{\hbar}{2}(\nabla_\alpha f) g^{\alpha\beta} (\nabla_\beta k) = \frac{\hbar}{2} g(X_f, X_k). \tag{3.44}$$

Now, suppose that $f, k : \mathcal{P} \to \mathbb{R}$ are the projections of the expectation value functions $F, K : \mathcal{H} \to \mathbb{R}$. Let

$$H = \{F, K\}_+ \tag{3.45}$$

be the Riemann bracket of F and K and denote by h the corresponding function on $\mathcal{P}$. We seek the functional relationship between f, k and h, which is induced by Eq. (3.45). This is easily obtained by working with the lifts. From Eq. (3.21), we must have $\tilde{h} = \frac{2\hbar}{r^2}\{F, K\}_+$, which, by use of Eqs. (3.19) and (3.33), gives

$$\begin{aligned} \tilde{h} &= \frac{\hbar^2}{r^2} G(X_F, X_K) = \frac{\hbar}{2} G' \left(\tilde{X}_f + \frac{\tilde{f}}{\hbar}\mathcal{J},\ \tilde{X}_k + \frac{\tilde{f}}{\hbar}\mathcal{J} \right) \\ &= \frac{\hbar}{2} \tilde{g}(\tilde{X}_f, \tilde{X}_k) + \tilde{f}\tilde{k}. \end{aligned} \tag{3.46}$$

By projecting this expression to $\mathcal{P}$, we obtain

Lemma 3.6. *Let $F, K : \mathcal{H} \to \mathbb{R}$ be two expectation value functions and suppose $f, k :\to \mathbb{R}$ are the corresponding functions on the projective space. Then*

$$\{F,\ K\}_+|_S = \pi^*[(f,\ k) + fk].$$

We will therefore utilize another notational abuse and write

$$\{f,\ k\}_+ := (f,\ k) + fk, \tag{3.47}$$

which we shall call the *symmetric bracket* of f and k.

The Riemann bracket of f and k does not correspond precisely to the Jordan product of the corresponding observables; in fact (f, k) is, in general, not even an observable function. It does, however, have an immediate physical interpretation. Recall that the restriction of $\tilde{f}$ to S is simply the expectation $F|_S$. An immediate consequence of Eq. (2.31) is the expression for the uncertainty of $\hat{F}$ at the state $p \in \mathcal{P}$:

$$(\Delta f)^2(p) := (\Delta \hat{F})^2(\pi^{-1}(p)) = (f, f)(p). \tag{3.48}$$

The second expression is, of course, well-defined, since $\Delta \hat{F}$ is constant on $\pi^{-1}(p)$. The uncertainty of an observable is then naturally expressed in terms of the Riemann bracket.

As a special case, we can recover a result obtained by Anandan and Aharonov [12]. Suppose that the Hamiltonian operator is bounded and consider the uncertainty in the energy. Let h be the projection, to $\mathcal{P}$, of the expectation value of the Hamiltonian operator. From Eq. (3.44), $(\Delta h)^2 = \frac{\hbar}{2} g(X_h, X_h)$. Therefore, apart from the constant coefficient, the uncertainty in the energy is exactly the norm of the Hamiltonian vector field which generates the time-evolution. The energy uncertainty is essentially the 'speed with which the system moves through the quantum phase space'.

Next, let $f, k : \mathcal{P} \to \mathbb{R}$ correspond to the expectation value functions F and K, as in Lemma 3.6. Comparing Eqs. (2.36) and (3.47), we have

$$(f,\, k)(p) = \{f,\, k\}_+(p) - f(p)k(p) = \mathrm{Cov}_{\pi^{-1}(p)}(\hat{F}, \hat{K}), \tag{3.49}$$

where the last expression is interpreted as in Eq. (3.48). Thus, it is the quantum covariance that is naturally expressed in terms of the Riemann bracket defined by g. In particular, the uncertainty relation written in Eq. (2.35) takes the form

$$(\Delta f)^2(\Delta k)^2 \geq \left(\frac{\hbar}{2}\{f,\, k\}\right)^2 + (f, k)^2. \tag{3.50}$$

Once again, if the standard Heisenberg uncertainty relation,

$$(\Delta f)^2(\Delta k)^2 \geq \left(\frac{\hbar}{2}\{f,\, k\}\right)^2, \tag{3.51}$$

is saturated at $p \in \mathcal{P}$, then p is a state of minimal 'coherence' in the sense that the quantum covariance of f and k vanishes at p. In fact, as a direct result of Shankar's presentation [27, Ch. 9], the Heisenberg uncertainty relation, expressed in Eq. (3.51), is saturated at p if and only if $(f, k)(p) = 0$ *and* $X_f(p) \propto {}_J X_k(p)$.

3.2.3 The measurement process

The issues that arise in the description of the measurement process are not only peculiar to quantum mechanics, but are at the very heart of quantum mechanics. It may be said that one of the primary purposes of the quantum mechanical formalism is to model the measurement process. If our formalism were unable to incorporate an intrinsic description of transition amplitudes, the program would

be doomed. Fortunately, this is not the case. We now treat the geometric meaning of transition probabilities and the associated "collapse of the wave function".

Transition probabilities

Let Ψ be an arbitrary element of the Hilbert space with unit normalization, $\langle \Psi, \Psi \rangle = 1$. Of obvious interest, in the context of measurement, is the function $\tilde{P}_\Psi(\Phi) := |\langle \Psi, \Phi \rangle|^2$, which we view as a function on S; i.e., we consider only $\langle \Phi, \Phi \rangle = 1$. Since $\tilde{P}_\Psi(\Phi)$ is independent of the phase of both Ψ and Φ, it determines a two-point function $P_q(x)$ on $\mathcal{P}$, defined as

$$P_q(x) = \tilde{P}_{\pi^{-1}q}(\pi^{-1}x) = \left| \left\langle \pi^{-1}(q),\, \pi^{-1}(x) \right\rangle \right|^2 . \tag{3.52}$$

We would like to know how to calculate this function in a way intrinsic to the projective space.

Fix $q \in \mathcal{P}$ arbitrarily, and let Ψ be an arbitrary element of $\pi^{-1}(q)$. Similarly, given $x \in \mathcal{P}$, choose an element $\Phi \in \pi^{-1}(x)$. Without loss of generality, we may assume that $\langle \Psi, \Phi \rangle$ is real and non-negative; otherwise, simply make the replacement $\Phi \mapsto \frac{\langle \Phi, \Psi \rangle}{|\langle \Phi, \Psi \rangle|}\Phi$. This is only ill-defined if $\langle \Phi, \Psi \rangle = 0$, in which case we have no need to rotate Φ in the first place. Note that, given Ψ, the above procedure uniquely determines Φ, unless $\langle \Psi, \Phi \rangle = 0$. For convenience, let us define a third element of S via

$$\Psi_1 := \left[1 - \langle \Psi,\, \Phi \rangle\right]^{\frac{1}{2}} \left(\Phi - \langle \Psi,\, \Phi \rangle \Psi\right). \tag{3.53}$$

By construction, $\langle \Psi, \Psi_1 \rangle = 0$, since Ψ_1 is just obtained by the Gram-Schmidt process. Finally, consider the curve

$$\begin{aligned} \eta(t) &:= \Psi \cos(t) + \Psi_1 \sin(t) \\ &= \left(\cos(t) - \frac{\langle \Psi, \Phi \rangle \sin(t)}{\left[1 - |\langle \Psi, \Phi \rangle|^2\right]^{\frac{1}{2}}} \right) \Psi + \frac{\sin(t)}{\left[1 - |\langle \Psi, \Phi \rangle|^2\right]^{\frac{1}{2}}} \Phi \end{aligned} \tag{3.54}$$

For future reference, it will be useful to note that this curve passes through Φ at the parameter value $t = \cos^{-1}(\langle \Psi, \Phi \rangle)$.

Of course, $\eta(t)$ defines a closed curve, $p(t) = [\Psi(t)]$ on the projective space. We will see that $p(t)$ is a geodesic; in fact,

Theorem 3.7. *Given arbitrary points $q, x \in \mathcal{P}$, there exists a (closed) geodesic, which passes through q and x. Further, $P_q(x) = \cos^2(\frac{\sigma(q,x)}{\sqrt{2\hbar}})$, where $\sigma(q,x)$ is the geodesic separation of q and x.*

Proof. We show that the curve $p(t)$, introduced above, is a geodesic. We need to show that $v^\beta \nabla_\beta v^\alpha = 0$, where v^α is the tangent to $p(t)$. Let us work in terms of the lift, $\eta(t)$. Define $\tilde{v}(t) := \dot{\eta}(t) = -\Psi \sin(t) + \Psi_1 \cos(t)$. Since $\eta(t) \in S$, $\tilde{v}^a(t)$ is G-orthogonal to the vector $\eta(t)$. It is also easy to see that $\mathcal{J}_a|_{\eta(t)}\, \tilde{v}^a(t) = 0$. Thus, as suggested by the notation, $\tilde{v}^a(t)$ is horizontal; $\eta(t)$ is a horizontal lift of $p(t)$, and $\tilde{v}^a(t)$ is the corresponding horizontal lift of $v^\alpha(t)$. Now, $\tilde{v}^b D_b \tilde{v}^a(t) = \ddot{\eta}^a(t) = -\eta^a(t)$. Thus, using Eqs. (3.31) and (3.36),

$$\tilde{v}^b \tilde{\nabla}_b \tilde{v}^a(t) = \tilde{v}^b \tilde{g}^a{}_d \left[D_b \tilde{v}^d + C^d{}_{bc} \tilde{v}^c \right](t) = \tilde{g}^a{}_d \left[\tilde{v}^b D_b \tilde{v}^d + \eta^a(t) \right] = 0.$$

Equivalently, $v^\beta \nabla_\beta v^\alpha = 0$. This proves the first statement.

Next, since $|\langle \dot\eta(t), \dot\eta(t)\rangle| = 1$, $g(v(t), v(t)) = \tilde g_{ab}|_{\eta(t)}\, \tilde v^a(t)\tilde v^b(t) = 2\hbar$. Thus, the length along $p(t)$ is $\int_0^t \sqrt{v^\alpha v_\alpha}\mathrm{d}t = t\sqrt{2\hbar}$. From Eq. (3.52) we then obtain $P_q(p(t)) = |\langle \Psi, \eta(t)\rangle|^2 = \cos^2(t) = \cos^2\left(\frac{\sigma(q,p(t))}{\sqrt{2\hbar}}\right)$, where $\sigma(q, p(t))$ is the geodesic length along $p(t)$. Choosing $t = \cos^{-1}(\langle \Psi, \Phi\rangle)$, gives the desired result. □

Incidentally, one can argue that if $\left|\langle \pi^{-1}q, \pi^{-1}x\rangle\right| \neq 0$, then the geodesic $p(t)$ above is unique, up to reparametrization; i.e., the *image* of $p(t)$ is the same for any geodesic connecting q and x. However, strictly speaking, the geodesic separation $\sigma(q, x)$ is ill-defined. For example, the direction along the curve connecting q and x, and the number of times one winds along this curve will affect the value of $\sigma(q, x)$. However, since the distance traveled in winding once around is $\pi\sqrt{2\hbar}$, the quantity $\cos^2\left(\frac{\sigma(q,p(t))}{\sqrt{2\hbar}}\right)$ is well-defined. Nonetheless, for the sake of precision, by $\sigma(q, x)$, we will mean the *minimal* geodesic distance separating q and x, which may have been one's interpretation in the first place.

Suppose one is dealing with the measurement of an observable $\hat F$ with discrete, non-degenerate spectrum. Each eigenspace of $\hat F$ is one (complex) dimensional, and therefore determines a single point of $\mathcal{P}$. Denote by p_i these eigenstates of $\hat F$. Now, if the system is in the state labeled by the point q, when an ideal measurement of $\hat F$ is performed, we know that the system will 'collapse' to one of the states p_i. Theorem 3.7 gives us the corresponding probabilities. It is interesting to notice that the probability of collapse to an eigenstate p_i is a monotonically decreasing function of the (minimal) geodesic distance separating q and p_i. The system is more likely to collapse to a nearby state than a distant one.

State reduction

We have a description of the probabilities associated with the measurement process. We now consider the measurement issue in more detail. We are interested in describing the measurement of an observable $\hat F$ at an instant in which the system is in the state $q \in \mathcal{P}$. For simplicity, let us assume that the spectrum if $\hat F$ is discrete; in this case, the only possible results of the measurement are eigenvalues of $\hat F$.

Since we wish to have a formalism which does not refer to the underlying Hilbert space, we need a characterization of the eigenstates of $\hat F$ which is intrinsic to $\mathcal{P}$. This is very simple. Let $F : \mathcal{H} \to \mathbb{R}$ and $f : \mathcal{P} \to \mathbb{R}$ be the expectation value and corresponding observable function, respectively. A vector $\Psi \in \mathcal{H}$ is an eigenstate of $\hat F$ iff $\hat F\Psi = \lambda\Psi$, for some (real) λ. Equivalently, by inspection of Eq. (2.24),

$$X_F|_\Psi = Y_{\hat F}\big|_\Psi = (\lambda/\hbar)\ \mathcal{J}|_\Psi\,.$$

Ψ is an eigenstate of $\hat F$ if and only if the Hamiltonian vector field, X_F, is *horizontal* at Ψ. By virtue of Eq. (3.33), this will be the case if and only if X_f vanishes at p; p is then a critical point of f. Evidently, the corresponding eigenvalue is exactly the (critical) value of f at p. We state this as

Definition 3.3. Let $f : \mathcal{P} \to \mathbb{R}$ be an observable function. Critical points of f are called *eigenstates* of f. The corresponding critical values are called *eigenvalues*.

The description of the measurement of an observable with discrete, non-degenerate spectrum now takes the following form. Immediately after measurement of f, the system, by some miraculous process, will have been be forced in an eigenstate p_i. "Realizing that it has collapsed" to the eigenstate p_i, the system returns what it knows to be the value of the observable under consideration, i.e., $f(p_i)$. The corresponding probabilities are determined by the geodesic distances which separate q and p_i, as written in Theorem 3.7. Just as in classical mechanics, it is the value of the observable function f that is obtained upon measurement. However, unlike classical mechanics, the measurement process disturbs the system, forcing it into an eigenstate before the measured value is produced.

Let us now relax the assumption of non-degeneracy of the spectrum of $\hat{F}$; we still assume it to be discrete. Let λ be a degenerate eigenvalue of $\hat{F}$ and $\tilde{\mathcal{E}}_\lambda$ be the associated eigenspace in $\mathcal{H}$. Associated to the this eigenspace is a submanifold, $\mathcal{E}_\lambda$, of $\mathcal{P}$, which we shall call the *eigenmanifold* associated to λ. Suppose the quantum system is prepared in the state q, and let $\Psi \in \pi^{-1}(q)$; i.e., $\langle \Psi, \Psi \rangle = 1$. The postulates of standard quantum mechanics state that the probability that measurement of f will yield the value λ is given by $\langle \mathbb{P}_\lambda(\Psi), \mathbb{P}_\lambda(\Psi) \rangle = \langle \Psi, \mathbb{P}_\lambda(\Psi) \rangle = \left| \left\langle \Psi, \frac{\mathbb{P}_\lambda(\Psi)}{\| \mathbb{P}_\lambda(\Psi) \|} \right\rangle \right|^2$. We know, from the above considerations, that this probability may be expressed in terms of the geodesic separation of $q = [\Psi]$ and the particular point $[\mathbb{P}_\lambda(\Psi)/\| \mathbb{P}_\lambda(\Psi) \|] \in \mathcal{E}_\lambda$. Let us denote by $\mathbb{P}_\lambda(q)$, this special element of the eigenmanifold.

What is it that picks out this point from all the other elements of $\mathcal{E}_\lambda$? In order to answer this question, we only need to notice that for any point $\Phi \in \tilde{\mathcal{E}}_\lambda \cap S$,

$$\begin{aligned} |\langle \Psi, \Phi \rangle|^2 &= |\langle \Psi, \mathbb{P}_\lambda(\Phi) \rangle|^2 = |\langle \mathbb{P}_\lambda(\Psi), \Phi \rangle|^2 \\ &\leq \| \mathbb{P}_\lambda(\Psi) \|^2. \end{aligned} \tag{3.55}$$

Therefore, of all elements $\Phi \in \tilde{\mathcal{E}}_\lambda$, with unit normalization, that which maximizes the quantity $|\langle \Psi, \Phi \rangle|^2$ is simply $\mathbb{P}_\lambda(\Psi)/\| \mathbb{P}_\lambda(\Psi) \|$, i.e., the one to which the state Ψ will collapse in the event that measurement of f yields the value λ. The implication of this observation is the following. For each point e $\in \mathcal{E}_\lambda$, we may calculate the geodesic separation $\sigma(q, e)$. The point of the eigenmanifold for which $\sigma(q, e)$ is minimal is, of course, that for which $\cos^2\left(\frac{\sigma(q,e)}{\sqrt{2\hbar}}\right)$ is maximized. $\mathbb{P}_\lambda(q)$ is that point of $\mathcal{E}_\lambda$ which is nearest q. Let us call $\sigma(q, \mathbb{P}_\lambda(q))$ the *geodesic separation of q and $\mathcal{E}_\lambda$*, or simply the distance between q and $\mathcal{E}_\lambda$. This distance will also be denoted $\sigma(q, \mathcal{E}_\lambda)$.

Regarding the measurement of an observable with discrete spectrum, we may now say the following. Suppose that immediately prior to measurement of f, the system is in the state $q \in \mathcal{P}$. Denote by λ_i the critical values of f and, by $\mathcal{E}_{\lambda_i}$, the corresponding eigenmanifolds. Interaction with the measuring device causes the system to be projected to one of the eigenmanifolds. Suppose the lucky eigenmanifold is $\mathcal{E}_{\lambda_0}$; then the system is found in the state $\mathbb{P}_{\lambda_0}(q)$—that point of $\mathcal{E}_{\lambda_0}$ which is nearest q. The probability of $\mathcal{E}_{\lambda_0}$ being the 'chosen one' is simply given by the geodesic separation of q and $\mathcal{E}_{\lambda_0}$, by replacing x with $\mathcal{E}_{\lambda_0}$ in Theorem 3.7. The value obtained, as a result of this measurement,

is then the value of f at the projected state.

Finally, let us study the generic situation. Let $\hat{F}$ be any observable of the theory, the spectrum of which is allowed to be continuous. As a first step, we need a definition of the spectrum of an observable $\hat{F}$ in terms of the corresponding function $f : \mathcal{P} \to \mathbb{R}$. Recall the standard definition: the (real) number λ is an element of the spectrum of $\hat{F}$ if $\hat{F} - \lambda \mathbf{1}$ is *not* invertible. This definition stems from the intuitive statement that such a λ is "nearly an eigenvalue". An alternative definition, which makes contact with this intuitive idea is the following: λ is an element of the spectrum of $\hat{F}$ if given any $\varepsilon \in \mathbb{R}$, $\exists\, \Psi \in S$ such that $\| \hat{F}\Psi - \lambda\Psi \| < \varepsilon^2$. This condition states that there exists a (normalized) state which is as close as one wishes to being an eigenstate of $\hat{F}$. Recalling the expressions (3.47) and (3.48), for the Jordan product and uncertainty, respectively, the (square of the) above quantity may be written as

$$\begin{aligned} \| \hat{F}\Psi - \lambda\Psi \|^2 &= \langle \Psi,\, (\hat{F} - \lambda)^2 \Psi \rangle \\ &= \{ f - \lambda, f - \lambda \}_+ \big|_{\pi(\Psi)} \\ &= \left[(\Delta f)^2 + (f - \lambda)^2 \right] \Big|_{\pi(\Psi)} . \end{aligned} \tag{3.56}$$

Therefore, one may define the spectrum of the observable function f as follows. A real number λ is an element of the spectrum of the observable function f if for any ε $\exists\, p \in \mathcal{P}$ such that $(\Delta f|_p)^2 + (f|_p - \lambda)^2 < \varepsilon^2$. The state p is one at which f takes on a value close to λ *and* for which the uncertainty of f is small. It is the existence of a sequence of wave packets (approximate eigenstates) at which the uncertainty of f tends to zero while the value of f tends to λ which guarantees that λ be an element of the spectrum of f. For the position operator, for example, one should have in mind a sequence of functions which tend to the 'Dirac delta function'. Let us state the result formally as

Definition 3.4. The *spectrum* $\mathrm{sp}(f)$ of an observable f consists of all real numbers λ for which the function $n_\lambda : \mathcal{P} \to \mathbb{R} \cup \{\infty\}$, $n_\lambda : p \mapsto \left[\big(\Delta f\big)^2(p) + \big(f(p) - \lambda\big)^2 \right]^{-1}$ is *unbounded*.

Of course, a point at which n_λ actually 'blows up' corresponds to an eigenstate of f.

The next step is a description of the spectral projection operators. Let Λ be a closed subset of the spectrum $\mathrm{sp}(f)$ of f. Associated to Λ is the projection operator $\mathbb{P}_{\hat{F},\Lambda}$, as guaranteed by the spectral theorem.[5] In analogy with what was done for the discrete case, we define

$$\tilde{\mathcal{E}}_{\hat{F},\Lambda} := \left\{ \mathbb{P}_{\hat{F},\Lambda}(\Psi) \,|\, \Psi \in \mathcal{H} \right\}.$$

The set $\tilde{\mathcal{E}}_{\hat{F},\Lambda}$—the image of $\mathbb{P}_{\hat{F},\Lambda}$—is the analog of the eigenspace of the previous discussion. In fact, it *is* the eigenspace of $\mathbb{P}_{\hat{F},\Lambda}$ corresponding to the eigenvalue 1. Now let $\mathcal{E}_{f,\Lambda}$ denote the projection of $\tilde{\mathcal{E}}_{\hat{F},\Lambda}$ to $\mathcal{P}$. We would like to characterize $\mathcal{E}_{f,\Lambda}$ in a way which is intrinsic to $\mathcal{P}$.

First, let us assume that Λ contains only non-negative elements. This assumption, which may easily be removed, only serves to eliminate a minor technicality that will become apparent shortly.

[5] The projection operators introduced in spectral theory actually correspond to what we would write as $\mathbb{P}_{\hat{F},I_\lambda}$, where $I_\lambda = \mathrm{sp}(f) \cap (-\infty, \lambda]$. Our discussion is intended to add a slight simplification to the subject; at the end of the day, one typically constructs our projection operators from the spectral projections anyway.

It seems clear that $\Psi \in S$ is an element of $\tilde{\mathcal{E}}_{\hat{F},\Lambda}$ if and only if for every $n \in \mathbb{Z}^+$, $\langle \hat{F}^n \rangle_\Psi \in \Lambda^n$, where Λ^n denotes the image of Λ under the map $\lambda \mapsto \lambda^n$. (If Λ intersects the negative half-axis, this statement must be modified, but only slightly.) One should be able to see this by working in the representation of $\mathcal{H}$ which is defined by $\hat{F}$; we shall not, however, give a formal proof. To make contact with the case in which the spectrum of f is discrete, suppose Λ consists of the single eigenvalue λ. Then $\tilde{\mathcal{E}}_{\hat{F},\Lambda}$ must be the corresponding eigenspace; $\Psi \in \tilde{\mathcal{E}}_{\hat{F},\Lambda} \Leftrightarrow \hat{F}\Psi = \lambda\Psi$. As a trivial consequence $\langle \hat{F}^n \rangle_\Psi = \lambda^n \; \forall n \in \mathbb{Z}^+$. Similarly, one can verify that our statement is valid for the case in which Λ consists of a set of positive eigenvalues.

It is now a simple matter to define the projected 'eigenspace' $\mathcal{E}_{f,\Lambda}$. Recall the meaning of the symmetric bracket defined by Eq. (3.47). If $\hat{F}$ and $\hat{K}$ are two bounded, self-adjoint operators on $\mathcal{H}$ and f and k are the corresponding observable functions on $\mathcal{P}$ (i.e., the expectation value functions), then $\{f,k\}_+$ is the observable function which corresponds to (half) the anti-commutator of $\hat{F}$ and $\hat{K}$. Choosing $\hat{F} = \hat{K}$, we see that $\{f,f\}_+$ is simply the (projection to $\mathcal{P}$ of the) expectation value of $\hat{F}^2$. More generally, the expectation value of $\hat{F}^n$ projects to the n-fold symmetric product $\{f,\{f,\{f,\ldots\}_+\}_+\}_+$. Therefore, a state

$$q \in \mathcal{P} \text{ is an element of } \mathcal{E}_{f,\Lambda} \quad \Leftrightarrow \quad \{f,\{f,\{f,\overset{n \text{ terms}}{\ldots}\}_+\}_+\}_+\big|_q \in \Lambda^n \; \forall \; n \in \mathbb{Z}^+.$$

Having chosen a closed subset Λ of the spectrum of f and defined the corresponding 'eigenmanifold' $\mathcal{E}_{f,\Lambda}$, the definition of the corresponding projection operator is straight-forward. Let Ψ be an arbitrary element of $\mathcal{H}^\times$, put $q = \Pi(\Psi)$ and let $\mathbb{P}_{f,\Lambda}(q) := \Pi \circ \mathbb{P}_{\hat{F},\Lambda}(\Psi)$. One can now simply repeat the argument surrounding Eq. (3.55) to see that $\mathbb{P}_{f,\Lambda}(q)$ is that element of $\mathcal{E}_{f,\Lambda}$ which is *nearest* q. The measurement process may then be described as follows. Suppose the system is in the state $q \in \mathcal{P}$ when an experimenter decides to measure the observable f. Following the rules of quantum mechanics, she "asks the system" whether the value of f lies in Λ—some closed subset of $\mathrm{sp}(f)$ which, theoretically, she is free to choose. The experimental apparatus drives the system to one of two states—either $\mathbb{P}_{f,\Lambda}(q)$ or $\mathbb{P}_{f,\Lambda^c}(q)$, where Λ^c is (the closure of) the complement, in $\mathrm{sp}(f)$, of Λ. The system will be reduced the the former with probability

$$\cos^2\left(\frac{\sigma(q,\,\mathbb{P}_{f,\Lambda}(q))}{\sqrt{2\hbar}}\right); \tag{3.57}$$

in this event, the experiment yields the positive result ($f \in \Lambda$). The probability of collapse to the latter state is obtained by replacing Λ by Λ^c in Eq. (3.57), and in this case a negative result will be obtained. Of course, these probabilities sum to one (that this is not obvious from the construction is a disappointing feature of the formalism).

The above discussion of the spectral projections may seem a bit complicated and not quite 'natural'. It should be emphasised, however, that the definition of the spectral projection operators on the Hilbert space is not a trivial matter either. Spectral analysis is generally very far from simple.

Notice that our presentation reverses the typical description of the measurement process. Textbooks usually tell us how to calculate the transition probabilities, and then describe the reduced state corresponding to a given measured value. We have reversed the logical order in this discussion

to emphasize the following intriguing possibility for the description of the measurement process. The quantum system, when not subject to external influence, evolves along a Hamiltonian vector field, as in classical mechanics (a finite-dimensional picture is employed here). The interaction with a measurement device forces the system to evolve, not along a Hamiltonian trajectory, but along a geodesic. This geodesic evolution terminates after it has served its strange purpose of forcing the 'reduction'. The Hamiltonian evolution then continues. A model of the measurement process along these lines would be especially interesting in that the so-called reduction process would be a sort of dynamics which is described by the metric, which is, in general lacking in the classical mechanics. Such a model would further illuminate the purely quantum aspects of quantum mechanics, and possibly lead to implementations of Penrose's idea for gravitationally induced state-vector reduction (see the Introduction). Unfortunately, given an initial state and an observable, quantum mechanics only picks out a preferred class of geodesic curves, and associates corresponding probabilities to each of them. Another approach to the measurement problem, which also relies on the projective language we have been developing, is presented in a fascinating paper by Hughston [40].

3.3 Summary

We have successfully formulated quantum mechanics in a language which is intrinsic to the true space of quantum states—the projective Hilbert space $\mathcal{P}$. This state space is a Kähler manifold. As in classical mechanics, observables are real-valued functions which preserve the kinematic structure (though, in classical mechanics, any function satisfies this requirement). Being a Kähler manifold, $\mathcal{P}$ is, in particular, a symplectic manifold; the quantum symplectic structure serves precisely the purpose it ought to, from the classical point of view. The Poisson bracket not only preserves the set of observables, but exactly represents the familiar Lie bracket, determined by the commutator of linear operators. Further, if the Hamiltonian is bounded (as in the finite-dimensional case, for example), the Schrödinger evolution is described in the classical language of Hamiltonian mechanics (otherwise, as in classical mechanics, the evolution is not generated by an observable).

There are two features of quantum mechanics which are not present in the classical description. First, the phase space is of a very particular nature; it is a Kähler manifold of constant holomorphic sectional curvature (recall Eq. (3.39) and the remark which follows it). The second fundamental distinction between the classical and quantum formalisms is the presence, in quantum mechanics, of a Riemannian metric. While the symplectic structure serves exactly the same role as that of classical mechanics, the metric describes those features of quantum mechanics which do not have classical analogues. In particular, the uncertainty of an observable is essentially the length of its Hamiltonian vector field. The probabilistic features of quantum mechanics, as well as the so-called reduction process, have been seen to be very nicely described by the metric. It is this separation of 'U-related' structure verses 'R-related' structure which might be viewed as the most attractive feature of the description. This characterization of the two dynamical processes suggests, in particular, novel approaches to the measurement problem and to the study of the classical limit. The latter will be briefly considered in Ch. 5.

In a sense, one may now forget about the Hilbert space; the concepts involved in quantum mechanics may be described in a geometric language which refers to the projective space. Just as the inertial reference frames of special relativity are unnecessary fiducial structures, the Hilbert space may be viewed as an afterthought, with no *essential* purpose for the description of the physics. The linearity of quantum mechanics may be viewed as an artifact of a helpful structure, which is not really intrinsic to the theory. This is most easily summarized by restating the postulates in the geometric language.

3.3.1 A second look at the postulates

($\mathcal{H}$) *Physical states:* Physical states of the quantum system are in one-to-one correspondence with points of a Kähler manifold $\mathcal{P}$, which is necessarily a projective Hilbert space.

($\mathcal{U}$) *Kähler evolution:* The evolution of the system, when not subject to external influence, is determined by a flow on $\mathcal{P}$, which preserves the Kähler structure. In the finite-dimensional case, this flow is determined by integrating the Hamiltonian vector field of a preferred observable.

($\mathcal{O}$) *Observables:* Every measurable physical quantity is represented by a real-valued, smooth function on $\mathcal{P}$. Only those functions which preserve the Kähler structure are observable.

($\mathcal{P}$) *Probabilistic interpretation:* Let $\Lambda \subset \mathbb{R}$ be a closed subset of the spectrum of an observable f, and suppose the system is in the state corresponding to the point $p \in \mathcal{P}$. The probability that measurement of f will yield an element of Λ is given by

$$P_p(\Lambda) = \cos^2\left(\frac{\sigma(p,\ \mathbb{P}_{f,\Lambda}(p))}{\sqrt{2\hbar}}\right), \tag{3.58}$$

where $\mathbb{P}_{f,\Lambda}(p)$ is defined as above Eq.(3.57).

($\mathcal{R}_D$) *Reduction, discrete spectrum:* Suppose the spectrum of an observable f is discrete. This spectrum provides the set of possible outcomes of the ideal measurement of f. If measurement of f yields the eigenvalue λ, the state of the system immediately after the measurement is given by the associated projection, $\mathbb{P}_{f,\lambda}(p)$, of the initial state p.

($\mathcal{R}_C$) *Reduction, continuous spectrum:* A closed subset Λ of the spectrum of f determines an ideal measurement that may be performed on the system. This measurement corresponds to inquiring whether the value of f lies in Λ. Immediately after this measurement, the state of the system is given by $\mathbb{P}_{f,\Lambda}(p)$ or $\mathbb{P}_{f,\Lambda^c}(p)$, depending on whether the result of the measurement is positive or negative, respectively.

As in Ch. 2, Λ^c is the complement, in the spectrum of f, of the set Λ. Although the first 'reduction postulate' is a special case of the second, both have been included for comparison with the statements made in Ch. 2.

3.3.2 Example: The spin-$\frac{1}{2}$ system

Let us briefly consider the simplest of all possible examples—the spin-$\frac{1}{2}$ system. The Hilbert space is simply $\mathbb{C}^2$. The constraint space, S, is a 3-sphere of radius $\sqrt{2\hbar}$, and $\mathcal{P} = \mathbb{CP}^1$ is a 2-sphere. Recall that the geodesics on $\mathcal{P}$ come back to themselves after traveling a distance $\pi\sqrt{2\hbar}$. With our conventions, $\mathcal{P}$ is then a sphere of radius $\sqrt{\hbar/2}$.

Let us embed $\mathcal{P}$ in $\mathbb{R}^3$. The metric g is the restriction to $\mathcal{P}$ of a flat metric, h on $\mathbb{R}^3$;

$$\mathcal{P} = \{x \in \mathbb{R}^3 \mid h(x,x) = \hbar/2\}. \tag{3.59}$$

Up to re-scaling, there is only one symplectic structure on a 2-sphere; $\omega_{\alpha\beta}$ must be a multiple of $\varepsilon_{\alpha\beta\gamma}x^\gamma$, where ε is the alternating tensor defined by the metric, and x^γ is the radial vector field, defined as we have done with Ψ^a; i.e., a point $x \in \mathbb{R}^3$ is a vector, which may be identified with a tangent vector in $T_x\mathbb{R}^3$. By requiring that g and ω be compatible, one finds that the proportionality constant in unity;

$$\omega_{\alpha\beta}(x) = \varepsilon_{\alpha\beta\gamma}x^\gamma. \tag{3.60}$$

The observables are most easily described as follows. Let $\mathbf{v}$ be any point in $\mathbb{R}^3$, and consider the function $f_{\mathbf{v}}(x) := v_\alpha x^\alpha$, where the contraction is defined by h. Note that $X_f{}^\alpha = \omega^{\alpha\beta}\partial_b f_v = \varepsilon^{\alpha\beta}{}_\gamma v_\beta x^\gamma = (\mathbf{v}\times\mathbf{x})^\alpha$, where this last expression denotes the vector cross-product. The motion generated by $f_{\mathbf{v}}$ is, therefore, a rigid rotation of the sphere, whose fixed points are $\mathbf{v}$ and its antipode. These rigid rotations preserve g and ω; the $f_{\mathbf{v}}$ are observables. There are only three independent rotations of the 2-sphere so we are missing one observable (the number of linearly independent Hermitian matrices, and hence the dimension of the algebra of observable functions, is 4). The missing observables are simply the constant functions on $\mathcal{P}$. The generic quantum observable is then of the form

$$f(x) = f_{\mathbf{v}}(x) + v_0 = \mathbf{v}\cdot\mathbf{x} + v_0; \tag{3.61}$$

i.e., a combination of the first two spherical harmonics. One can make closer contact with the standard picture by choosing an orthonormal basis $\{e_{\underline{i}}\}$ for $\mathbb{R}^3$. The function $f_{\mathbf{v}}$ is then the projection, to $\mathcal{P}$, of the expectation value of $v^{\underline{i}}\hat{S}_{\underline{i}} = \frac{\hbar}{2}v^{\underline{i}}\hat{\sigma}_{\underline{i}}$, where $\hat{\sigma}_{\underline{i}}$ are the Pauli matrices.

Now let our spin-$\frac{1}{2}$ particle be influenced by a magnetic field $\mathbf{B}$. The Hamiltonian, in the operator language, is $\hat{H} = \gamma\mathbf{B}\cdot\mathbf{S}$; the corresponding observable function is, therefore, given by

$$h(x) = \gamma B_\alpha x^\alpha = \gamma f_{\mathbf{B}}(x). \tag{3.62}$$

Here, γ is the constant $\gamma = -\frac{e}{mc}$, where e and m are the electric charge and mass of our particle. The Hamiltonian evolution is, as above, a rigid rotation of $\mathcal{P}$, about the axis defined by the magnetic field B^α. The evolution is periodic with frequency γB, where B is the length, with respect to h, of B^α.

Of course, the eigenstates of the observable $f_{\mathbf{v}} + v_0$ are simply those points on the axis defined by $\mathbf{v}$; i.e. the points $\pm\frac{\hbar}{2}\mathbf{v}/|\mathbf{v}|$. The corresponding eigenvalues are $v_0 \pm |\mathbf{v}|$. The geodesic distance

between two points x_1 and x_2 of $\mathcal{P}$ is $\frac{\hbar}{2}\theta_{12}$, where θ_{12} is the angle between x_1 and x_2. It is then very easy to see that if the system is in the state x, the probability that measurement of $f_{\mathbf{v}}$ will yield the eigenvalue $\pm|\mathbf{v}|$ is $\frac{1}{2}+\frac{1}{2}\cos(\theta)$, where θ is the angle separating x and the associated eigenstate.

Finally, let us consider the algebraic structure on the space of observables. The Poisson bracket is given by

$$\begin{aligned}\{f_{\mathbf{v}}+v_0,\ f_{\mathbf{w}}+w_0\}(x) &= \{f_{\mathbf{v}},\ f_{\mathbf{w}}\}(x) = (\partial_\alpha f_{\mathbf{v}})\omega^{\alpha\beta}(\partial_\beta f_{\mathbf{w}})\,(x)\\ &= v_\alpha \varepsilon^{\alpha\beta}{}_\gamma x^\gamma w_\beta = (\mathbf{v}\times\mathbf{w})\cdot\mathbf{x} = f_{\mathbf{v}\times\mathbf{w}}(x).\end{aligned} \tag{3.63}$$

This is essentially the familiar expression for the commutator between the Pauli matrices. The quantum covariance is

$$\begin{aligned}(f_{\mathbf{v}}+v_0,\ f_{\mathbf{w}}+w_0)(x) &= (f_{\mathbf{v}},\ f_{\mathbf{w}})(x) = \frac{\hbar}{2}(\partial_\alpha f_{\mathbf{v}})g^{\alpha\beta}(\partial_\beta f_{\mathbf{w}})\,(x)\\ &= \frac{\hbar}{2}\left[h^{\alpha\beta}-\frac{2}{\hbar}x^\alpha x^\beta\right]v_\alpha w_\beta = \frac{\hbar}{2}\mathbf{v}\cdot\mathbf{w}-(\mathbf{v}\cdot\mathbf{x})(\mathbf{w}\cdot\mathbf{x}).\end{aligned} \tag{3.64}$$

Again, this is *not* an observable. Note that if $\mathbf{v}$ and $\mathbf{w}$ are orthogonal, then they are not correlated at the states $\pm\frac{\hbar}{2}(\mathbf{v}\times\mathbf{w})/|\mathbf{v}\times\mathbf{w}|$. Expressing the Jordan product in terms of the above,

$$\begin{aligned}\{f_{\mathbf{v}}+v_0,\ f_{\mathbf{w}}+w_0\}_+(x) &= (f_{\mathbf{v}},f_{\mathbf{w}})\,(x)+(f_{\mathbf{v}}+v_0)(f_{\mathbf{w}}+w_0)(x)\\ &= v_0 f_{\mathbf{w}}(x)+w_0 f_{\mathbf{v}}(x)+v_0w_0+\frac{\hbar}{2}\mathbf{v}\cdot\mathbf{w}\\ &= f_{v_0\mathbf{w}+w_0\mathbf{v}}+\left(v_0w_0+\frac{\hbar}{2}\mathbf{v}\cdot\mathbf{w}\right).\end{aligned} \tag{3.65}$$

In summary, observables may be represented by elements of $\mathbb{R}^4$. Denote by $(\mathbf{v},v_0)$ and $(\mathbf{w},w_0)$ two generic observables. The algebraic operations of taking the Poisson bracket and Jordan product are represented by the operations:

$$\{(\mathbf{v},v_0),\ (\mathbf{w},w_0)\} \ \mapsto (\mathbf{v}\times\mathbf{w},\,0) \qquad \text{and} \tag{3.66}$$

$$\{(\mathbf{v},v_0),\ (\mathbf{w},w_0)\}_+ \mapsto (v_0\mathbf{w}+w_0\mathbf{v},\ v_0w_0+\mathbf{v}\cdot\mathbf{w}). \tag{3.67}$$

4

(Quantization)^2

The quantum phase space is just that—a phase space. This phase space, if stumbled upon "at random", would likely be mistaken for that of some classical theory. A physicist, having spent many years studying quantum mechanics, would likely ask "What is the corresponding quantum theory?" This physicist, noticing that the symplectic manifold under consideration admits a Kähler metric, might construct the associated Kähler polarization and employ the method of geometric quantization.[1] In particular, she would search for a $U(1)$-bundle over $\mathcal{P}$ with a connection compatible with the symplectic structure ω. She would view *all* smooth functions on $\mathcal{P}$ as classical observables; the functions which have direct quantum analogues (the 'good functions') would be those which preserve the Kähler polarization.

The description of quantum mechanics presented in the previous chapter defines a 'fake' classical theory, which is obtained by essentially forgetting about the metric and allowing, as observables, *all* functions on the phase space. Let us now consider what will result from quantization of this fake classical theory. We have not simply stumbled upon the phase space; we know something about its origin. In §3.1.3, we mentioned that the unit sphere S, in $\mathcal{H}$, is a principal $U(1)$-bundle over $\mathcal{P}$ which is naturally equipped with a connection. The horizontal subspace of $T_\Psi S$ consists of those tangent vectors which are $(i^*\Omega)$-orthogonal to the generator, $\mathcal{J}$, of phase rotations. That this notion of horizontality is consistent with the definition of a connection is trivial to verify. The connection form which corresponds to this connection is given by $\mathcal{A}_a = \frac{1}{2\hbar}\mathcal{J}_a$. Notice that, since $D_a\mathcal{J}_b = -\Omega_{ab}$, the curvature of $\mathcal{A}$ is given by

$$\mathrm{d}\mathcal{A} = -\frac{1}{\hbar}(i^*\Omega) = -\frac{1}{\hbar}\pi^*\omega.$$

Therefore, not only is the pre-quantization condition satisfied, but we have already been supplied with the pre-quantum bundle! Moreover, since any observable function generates a Killing vector field, the Lie derivative with respect to which then annihilates the complex structure, all observable functions will be deemed 'good' by the above physicist. The observable functions therefore have immediate (quantum)^2 analogues. Our fake classical theory not only *can* be quantized, it is *begging* to be quantized.

[1] A brief outline of the geometric approach to quantization is outlined in Appendix B.

Application of 'the' geometric quantization procedure to our pretend classical theory will produce yet another quantum theory (which one might quantize again (which one might...)). Considering the amount of structure which is present, one might expect the (quantum)^2 theory to be something of physical relevance. The first natural guess is that the quantization procedure may be stable; i.e., that we will simply obtain the quantum theory from which we started. The other reasonable possibility is an alternative description of second quantization. Remarkably, we will find that, in a very well-defined sense, *both* of these guesses are correct!

In §4.1, we quantize $(\mathcal{P}, \omega)$ by applying the geometric quantization procedure with the obvious Kähler polarization; let us call this "re-quantization" to avoid confusion. We will see that the space of pre-(re-)quantum states which satisfy the polarization condition is naturally isomorphic to the original Hilbert space $\mathcal{H}$. Further, the re-quantum operator associated to a good function f is precisely that operator which corresponds to f in the sense of Theorem 3.4.

The sense in which re-quantization naturally describes what is called second quantization is discussed in §4.2. We observe that there are actually *numerous* quantizations of our pretend classical theory. The Fock space is seen to be the sum of all possible quantizations.

The material presented in this chapter is fairly technical, and relies heavily on the ideas surveyed in Appendix A. Both appendices should be viewed as prologues to this chapter; familiarity with the concepts presented there is assumed.

4.1 Stability of the Quantization Procedure

Since the required input for 're-quantization' is naturally provided by the available geometric structures, one may expect quantization of $(\mathcal{P}, \omega)$ to be a fairly simple task. This is, in fact, the case. In §4.1.1, we construct the line bundle associated to the pre-quantum bundle $S \xrightarrow{\pi} \mathcal{P}$. We also introduce a convenient 'representation' of the sections of this line bundle (i.e., pre-(re-)quantum states) in terms of *functions* on S. This allows us to make contact with the original Hilbert space. In §4.1.2, we obtain the general solution of the polarization condition. We find a one-to-one correspondence between these solutions and elements of $\mathcal{H}$, which maps the re-quantized observables to the operators from which they were originally defined. In this way, we will see that the quantization procedure simply reproduces the original quantum theory.

4.1.1 Pre-(re-)quantization

As mentioned above, we have already constructed the pre-quantum bundle $S \xrightarrow{\pi} \mathcal{P}$ which is required for the geometric quantization procedure. To this principal bundle is associated a complex line bundle $E \xrightarrow{\pi_E} \mathcal{P}$. Recall that this line bundle may be viewed as the quotient of $S \times \mathbb{C}$ with respect to the equivalence relation

$$(\Psi, c) \sim (\Psi u^{-1}, uc) \quad \forall\, u \in U(1). \tag{4.1}$$

An element of E is an equivalence class $[\Psi, c] = [\Psi u^{-1}, uc]$ of $S \times \mathbb{C}$.

Since the pre-quantum bundle is non-trivial, a section does not determine a globally defined complex-valued function on $\mathcal{P}$. However, to any section $\xi : \mathcal{P} \to E$ is a canonically associated a function $\xi^\star$ on the *bundle manifold* S, defined by

$$\xi \circ \pi(\Psi) =: [\Psi, \xi^\star(\Psi)]. \tag{4.2}$$

For lack of a better term, let us call $\xi^\star$ the *principal function* associated to the section ξ. By definition, $[\Psi, \xi^\star(\Psi)] = [\Psi u,\, u^{-1}\xi^\star(\Psi)] = [\Psi u, \xi^\star(\Psi u)]$; therefore, the principal function $\xi^\star$ must satisfy the identity

$$\xi^\star(\Psi u) = u^{-1}\xi^\star(\Psi). \tag{4.3}$$

Conversely, any smooth function $\xi^\star : S \to \mathbb{C}$ which satisfies Eq. (4.3) defines the section determined by Eq. (4.2). There is then a one-to-one correspondence between sections of E (pre-quantum states) and complex-valued functions on S which are homogeneous of degree -1. Since we are attempting to make contact with the original Hilbert space, it will be useful to work in terms of the principal functions, rather than the sections.

Next, let f be *any* smooth, real-valued function on $\mathcal{P}$, i.e., one of the fake classical observables. To any such function is associated the pre-quantum operator $\hat{O}_f$, which acts on the space $\mathcal{H}_{pre}$ of sections of E.[2] Recall that this operator is defined by

$$\hat{O}_f\xi := -\mathrm{i}\hbar\mathcal{D}_{X_f}\xi + f\xi, \tag{4.4}$$

where $\mathcal{D}$ denotes the covariant derivative defined by the connection. Naturally, we would like to know how to model the action of $\hat{O}_f$ in terms of the principal functions; i.e., we would like an explicit expression for $\left(\hat{O}_f\xi\right)^\star$.

We make two preliminary observations. First, using the connection $\mathcal{A}$, we may construct the horizontal lift, $\tilde{X_f}$, of the Hamiltonian vector field X_f. Note that, in the event that f is an observable function, $\tilde{X_f}$ agrees with (the restriction to S of) what we have previously denoted by the same symbol. The covariant derivative of a section ξ of E is, as defined in the appendix, naturally expressed in terms of the associated principal function; in particular,

$$\mathcal{D}_{X_f}\xi = [\Psi, \pounds_{\tilde{X}_f}\xi^\star] \quad \Rightarrow \quad (\mathcal{D}_{X_f}\xi)^\star = \pounds_{\tilde{X}_f}\xi^\star. \tag{4.5}$$

Next, since f is trivially Lie algebra (of $U(1)$) valued, we may construct the corresponding *fundamental vector field* on S,

$$f^\sharp\big|_\Psi := \frac{\mathrm{d}}{\mathrm{d}t}\Psi\mathrm{e}^{-\mathrm{i}tf\circ\pi(\Psi)}\bigg|_{t=0}. \tag{4.6}$$

Notice that for any section ξ,

$$\begin{aligned} f^\sharp(\xi^\star)\big|_\Psi &= \frac{\mathrm{d}}{\mathrm{d}t}\,\xi^\star\left(\Psi\mathrm{e}^{-\mathrm{i}tf\circ\pi(\Psi)}\right)\bigg|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\left[\mathrm{e}^{(\mathrm{i}tf\circ\pi(\Psi))}\xi^\star(\Psi)\right]\bigg|_{t=0} \\ &= \mathrm{i}f\circ\pi(\Psi)\xi^\star(\Psi). \end{aligned} \tag{4.7}$$

[2] By $\mathcal{H}_{pre}$ we mean the space of *all* sections of the line bundle. We will not concern ourselves with the introduction of a Hermitian inner-product since, at the end of the day, we will have a natural choice.

Therefore, $(f\xi)^\star = -\mathrm{i}\pounds_{f^\sharp}\xi^\star$.

Combining this last equation with Eq. (4.5), we obtain the desired expression for the action of $\hat{O}_f$:

$$\left(\hat{O}_f\xi\right)^\star = -\mathrm{i}\hbar\left[\pounds_{\tilde{X}_f}\xi^\star + \frac{\mathrm{i}\tilde{f}}{\hbar}\xi^\star\right] =: -\mathrm{i}\hbar\pounds_{\hat{X}_f}\xi^\star, \tag{4.8}$$

where, for convenience, and to emphasize its role in quantization, we have defined $\hat{X}_f := \tilde{X}_f + \frac{1}{\hbar}f^\sharp$. It seems rather interesting that f preserves the Kähler polarization if it is an observable function, and that, in this case, the vector field $\hat{X}_f$ is precisely (the restriction to S of) the Schrödinger vector field associated to the corresponding operator $\hat{F}$ on $\mathcal{H}$.

Quite a bit of notation has just been introduced. The space $\mathcal{H}_{pre}$ of pre-(re-)quantum states are sections of E; we may represent this space by a set of functions on the principal bundle S. To each and every 'classical observable' is associated a vector field on S. This vector field determines an operator on $\mathcal{H}_{pre}$ which may be represented by Lie differentiation of the principal functions. It should be emphasized that this not specific to our particular example. One may use this language to describe geometric quantization of the arbitrary classical theory [41].

4.1.2 Re-quantization

We have carried out the essential features of pre-quantization of the fake classical theory. The next step in the quantization procedure is to examine the polarization condition on the pre-quantum wave functions. The physical sections are those which are preserved by the Kähler polarization. Let us now understand what this means for re-quantization.

Recall that a vector $V \in T_p^{\mathbb{C}}\mathcal{P}$ is an element of the polarization subspace P_p if it is an eigenvector of the complex structure with eigenvalue $+\mathrm{i}$; $\jmath^\alpha{}_\beta V^\beta = \mathrm{i}V^\alpha$. The generic vector *field*, which is everywhere an element of the polarization, is of the form $V^\alpha = v^\beta(g_\beta{}^\alpha - \mathrm{i}\omega_\beta{}^\alpha)$. The polarization condition on the physical sections may then be expressed as

$$(g_\beta{}^\alpha - \mathrm{i}\omega_\beta{}^\alpha)\,\mathcal{D}_\alpha\xi = 0. \tag{4.9}$$

To make closer contact with the original Hilbert space, we can extend any principal function $\xi^\star$ to all of $\mathcal{H}$. We require that the extension, denoted by $\xi^\star_{ext}$, satisfy the requirement

$$\xi^\star_{ext}(c\Psi) = \bar{c}\,\xi^\star_{ext}(\Psi), \tag{4.10}$$

which is clearly consistent with Eq. (4.3) and defines a unique extension of $\xi^\star$ to all of $\mathcal{H}$. Such extensions will be smooth on $\mathcal{H}^\times$, but will not, in general, be differentiable at the origin. Conversely, any such function defines, by restriction to S, a section of E, and, hence, a pre-quantum state. The polarization condition, as stated in Eq. (4.9), is equivalent to the 'lifted' equation for the extended functions:

$$(\tilde{g}_b{}^a - \mathrm{i}\tilde{\omega}_b{}^a)\,D_a\xi^\star_{ext} = 0, \tag{4.11}$$

where we have implicitly used the fact that the covariant derivative of a section lifts to the Lie derivative of the principal function, as in Eq. (4.5). Of course, Eq. (4.11) is valid only on $\mathcal{H}^\times$.

Now, the derivatives of $\xi^\star_{ext}$ with respect to the generators of the rays of $\mathcal{H}$ are given by

$$\Psi^a D_a \xi^\star_{ext} = \xi^\star_{ext}, \tag{4.12}$$

$$\mathcal{J}^a D_a \xi^\star_{ext} = \mathrm{i}\xi^\star_{ext}. \tag{4.13}$$

The polarization condition, when written in terms of the extended functions, then becomes (c.f. Eqs. (3.27) and (3.28))

$$\begin{aligned} 0 &= (\tilde{g}_b{}^a - \mathrm{i}\tilde{\omega}_b{}^a)\, D_a \xi^\star_{ext} \\ &= (G_b{}^a - \mathrm{i}\Omega_b{}^a)\, D_a \xi^\star_{ext} - \frac{1}{r^2}\left(\Psi_b \Psi^a + \mathcal{J}_b \mathcal{J}^a + \mathrm{i}\mathcal{J}_b \Psi^a - \mathrm{i}\Psi_b \mathcal{J}^a\right) D_a \xi^\star_{ext} \\ &= (G_b{}^a - \mathrm{i}\Omega_b{}^a)\, D_a \xi^\star_{ext}. \end{aligned} \tag{4.14}$$

We have arrived at the following: The space of pre-quantum states is isomorphic to the space of functions on $\mathcal{H}$ which are anti-homogeneous, in the sense of Eq. (4.10). Of these functions, those which represent physical states of the re-quantized theory are those which are preserved by the Kähler polarization *on the original Hilbert space*. Therefore, the 're-quantum' state space consists of the functions $\xi^\star_{ext} : \mathcal{H} \to \mathbb{C}$ which satisfy the following two equations:

$$\begin{aligned} &i) \qquad \xi^\star_{ext}(c\Psi) = \bar{c}\, \xi^\star_{ext}(\Psi), \\ &ii) \qquad (G_b{}^a - \mathrm{i}\Omega_b{}^a)\, D_a \xi^\star_{ext} = 0 \qquad (\text{on } \mathcal{H}^\times). \end{aligned} \tag{4.15}$$

One may recognize $\frac{1}{2}\left(G_a{}^b - \mathrm{i}\Omega_a{}^b\right) D_b$ as the Dolbeault operator ∂. Condition (ii) then states that $\xi^\star_{ext}$ is *anti-holomorphic* on $\mathcal{H}^\times$. Since holomorphic functions are fairly 'rigid', we should expect the conditions expressed in Eqs. (4.15) to be fairly strong.

In order to explain this, and to avoid infinite-dimensional complex analysis, let us pull back Eqs. (4.15) to an arbitrary finite-dimensional, complex subspace V of $\mathcal{H}$. That V is a complex subspace means only that it is preserved by the complex structure $J^a{}_b = \Omega_b{}^a$. Suppose the complex dimension of V is n, and let $\{\mathrm{e}_1, \ldots, \mathrm{e}_n\}$ be an orthonormal basis of V (with respect to the Hermitian inner-product $\langle , \rangle$). Then $\{\mathrm{e}_1, \ldots, \mathrm{e}_n, J\mathrm{e}_1, \ldots J\mathrm{e}_n\}$ is a real basis for V and any element $\Psi \in V$ may be uniquely written

$$\Psi = x^{\underline{i}} \mathrm{e}_{\underline{i}} + y^{\underline{i}} J\mathrm{e}_{\underline{i}}. \tag{4.16}$$

We are simply introducing complex coordinates $z^{\underline{i}} = x^{\underline{i}} + \mathrm{i}y^{\underline{i}}$ on V. (This is a rare instance in which our notation, which slightly disguises the complex nature of $\mathcal{H}$, is inconvenient.) Let ζ be any complex-valued function on $\mathcal{H}$ which satisfies conditions (i) and (ii) (and therefore represents a re-quantum state), and let $\zeta|_V$ denote its restriction to V. As a result of condition (ii), we find that $\zeta|_V$ satisfies

$$\left(\frac{\partial}{\partial x^{\underline{i}}} - \mathrm{i}\frac{\partial}{\partial y^{\underline{i}}}\right) \zeta|_V = 0 \qquad (\text{on } V - \{\mathbf{0}\}), \tag{4.17}$$

and is therefore anti-holomorphic on $V-\{\mathbf{0}\}$. Further, by virtue of condition (i), this function is bounded on any bounded region containing $\mathbf{0}$. Thus, by a standard result of complex analysis (see, for example, Rudin [42]), $\zeta|_V$ extends to an anti-holomorphic function on all of V. It's complex conjugate is therefore representable by a power series in $z^{\underline{i}}$ about $\mathbf{0}$, which by (i), must consist of only the linear terms. Hence $\zeta|_V$ is a continuous, anti-linear function of $z^{\underline{i}}$. Finally, since this is true for arbitrary V, ζ is a continous, anti-linear function on $\mathcal{H}$, and must therefore be of the form $\zeta(\Psi) = \langle \Psi,\, \Phi_\zeta \rangle$.

We have just seen that the general solution to Eqs. (4.15) is given by

$$\xi^\star_{ext}(\Psi) = \langle \Psi, \Phi_\xi \rangle. \tag{4.18}$$

The re-quantum Hilbert space is, in this way, naturally isomorphic to the original Hilbert space. In particular, there is a natural inner-product to choose for the 're-quantized' theory.

Finally, let us consider the re-quantized observables. We already know that the quantizable functions are precisely the observable functions. Let f be any such function on $\mathcal{P}$, and let $\hat{F}$ be the corresponding operator on $\mathcal{H}$. Since $\hat{X}_f$ is the Schrödinger vector field associated to $\hat{F}$, for any section ξ that satisfies the polarization condition, we have

$$\begin{aligned}
\left(\hat{O}_f \xi\right)^\star (\Psi) &= -\mathrm{i}\hbar \pounds_{\hat{X}_f} \xi^\star(\Psi) \\
&= -\mathrm{i}\hbar \lim_{\varepsilon \to 0} \left[\left\langle \Psi - \frac{\varepsilon}{\hbar} J\hat{F}\Psi,\, \Phi_\xi \right\rangle - \langle \Psi,\, \Phi_\xi \rangle \right] \\
&= \left\langle \hat{F}\Psi,\, \Phi_\xi \right\rangle \;=\; \left\langle \Psi,\, \hat{F}\Phi_\xi \right\rangle .
\end{aligned} \tag{4.19}$$

Therefore, under the mapping $\xi \mapsto \Phi_\xi$, the re-quantized operator $\hat{O}_f$ simply maps to the operator $\hat{F}$ from which f was defined in the first place!

Our first guess was correct; (quantization)^2 = quantization. We already know this result to hold for the simplest case—the spin-$\frac{1}{2}$ system. The classical model of spin is the 2-sphere, equipped with a symplectic structure. Quantization of this 'classical theory' simply produces the spin-$\frac{1}{2}$ Hilbert space $\mathbb{C}^2$ (that is, if one chooses the 'correct' symplectic structure; see below). As we know, this classical theory is exactly the fake one determined by the structure on the projective space $\mathbb{CP}^1$.

Actually, the story is slightly more involved than that. There is *one* symplectic structure, say ω_1 on S^2, quantization of which produces the Hilbert space $\mathbb{C}^2$. All other symplectic structures on S^2 which satisfy the pre-quantization condition are integral multiples of ω_1. The other finite-dimensional Hilbert spaces arise by quantization of these multiples of ω_1 (see, for example, Woodhouse [43]). In this way, the 2-sphere is the classical phase space corresponding to all finite-dimensional quantum theories. We have found that one may also view $\mathbb{CP}^n$ as describing the 'classical theory' corresponding to the quantum Hilbert space $\mathbb{C}^{n+1}$.

4.2 Second Quantization

A moment's reflection will lead the way to a similar description of second quantization. As just discussed, there are countably many symplectic structures on S^2 which satisfy the pre-quantization

condition—those whose integrals are integer multiples of $2\pi\hbar$. It turns out that the the first quantizable symplectic structure on $S^2 = \mathbb{CP}^1$ is naturally induced by the inner-product on $\mathcal{H} = \mathbb{C}^2$. Or is it?

The fact of the matter is that, from a mathematical point of view, there is noting 'natural' about our particular choice of symplectic structure on $\mathcal{P}$. We chose to define ω by projecting from the unit sphere in $\mathcal{H}$ because of our 'physical' biases, not because it is the mathematically preferred choice. Had we chosen some other sphere, say

$$S_k := \{\Psi \in \mathcal{H} \mid \langle \Psi, \Psi \rangle = k\}, \quad \text{any real } k, \tag{4.20}$$

the entire game could have been played with no lack of success. Had we done that, we would have obtained the symplectic structure $\omega_k := k\omega$, which is, in general, *not* quantizable. The fact that we just quantized $(\mathcal{P}, \omega)$ is clear enough indication that ω_k is quantizable, for $k = 1$. It then follows that ω_k satisfies the pre-quantization condition for all *integral* values of k. Furthermore, these are the only values of k for which ω_k is quantizable, since the integral of ω over the 2-sphere obtained by projecting any 2-complex dimensional subspace of $\mathcal{H}$ is exactly $2\pi\hbar$.

We will now quantize $(\mathcal{P}, \omega_k)$ for integer k. Recall the result for $\mathcal{H} = \mathbb{C}^2$: quantization of ω_k produces the spin-$\frac{k}{2}$ Hilbert space, $\mathbb{C}^{k+1}$. For this reason, one might guess that quantization of ω_k will yield the k-particle subspace of the second quantized theory. We will see that this guess is correct.

4.2.1 Pre-(second-)quantization

For the case $k = 1$ we had already constructed the pre-quantum bundle. In the general case, we are not so lucky. Now, we are imagining that we have described the geometric structure (g_k, ω_k) on $\mathcal{P}$ by projection from the sphere S_k. This sphere is also a principal bundle over $\mathcal{P}$ (it is essentially identical to $S \xrightarrow{\pi} \mathcal{P}$); and S_k is equipped with the natural (Berry) connection determined by the one-form $\frac{1}{2k\hbar}\mathcal{J}$. Unfortunately, the curvature of this connection is given by $-\frac{1}{k\hbar}{\pi_k}^*\omega_k$, where $\pi_k : S_k \to \mathcal{P}$ denotes the relevant projection. Note that multiplication, by k, of the above connection form does *not* produce another connection form on S_k. The construction of the pre-quantum bundle, in the general case, is not the trivial matter encountered in the previous section.

In order to construct the right pre-quantum bundle, we would like to "put some twists" in $S_k \xrightarrow{\pi_k} \mathcal{P}$. Since $U(1)$ acts on itself in a number of trivial ways, i.e.,

$$U(1) \ni g : u \mapsto g^k u, \quad \text{for any integer } k, \tag{4.21}$$

we can construct other $U(1)$ bundles, which are associated to this one. This is done as follows: Define equivalence relations on $S_k \times U(1)$ as

$$(\Psi, u) \approx_k (\Psi g^{-1}, g^k u) \quad \forall\, g \in U(1). \tag{4.22}$$

Denote the corresponding equivalence classes by $[\Psi, u]_k$, and the space of such equivalence classes by Σ_k. The result is a principal bundle, $\Sigma \xrightarrow{\tilde{\pi}_k} \mathcal{P}$, where the projection is defined as $\tilde{\pi}_k[\Psi, u]_k = \pi_k(\Psi)$.

The group action of $U(1)$ on this bundle is given by $[\Psi, u]_k \cdot g = [\Psi, gu]_k$. Note that the Ψ appearing in these expressions is an element of S_k, for some fixed k.

The bundle structure of $\Sigma_k \xrightarrow{\tilde{\pi}_k} \mathcal{P}$ is defined as follows. As for associated vector bundles, any local trivialization (U_i, ϕ_i) of S_k defines a corresponding trivialization $(U_i, \tilde{\phi}_i^k)$ of the associated principal bundle;

$$\tilde{\phi}_i^k : U_i \times U(1) \to \tilde{\pi}_k^{-1}(U_i) \qquad \tilde{\phi}_i^k : (p, u) \mapsto [\phi(p, e),\ u]_k. \tag{4.23}$$

However, unlike the situation discussed in Appendix A, for the case of associated *vector bundles*, the transition functions are *not* simply those of $S_k \xrightarrow{\pi_k} \mathcal{P}$ (after all, that's the point!). Let $(U_i, \tilde{\phi}_i^k)$ and $(U_j, \tilde{\phi}_j^k)$ be two local trivializations. Then, if g_{ij} is the transition function relating the trivializations of the original bundle,

$$\begin{aligned} \tilde{\phi}_j^k(p, u) &= [\phi_j(p, e),\ u]_k = [\phi_i(p, g_{ij}(p)),\ u]_k \\ &= [\phi_i(p, e) \cdot g_{ij}(p),\ u]_k = [\phi_i(p, e),\ g_{ij}^k(p) \cdot u]_k. \end{aligned} \tag{4.24}$$

So, to obtain the transition functions for the associated bundle, one simply raises the original transition functions to the k^{th} power.

Now, for each k, we have a new principal $U(1)$-bundle over $\mathcal{P}$. The next step is to define a connection $\mathcal{A}_k$ on each $\Sigma_k \xrightarrow{\tilde{\pi}_k} \mathcal{P}$, whose curvature is $-\frac{1}{\hbar}\omega_k$. There is an obvious candidate: the horizontal lifts of a curve $\gamma(t)$ on $\mathcal{P}$ are defined to be of the form $[\tilde{\gamma}(t),\ e]_k$, where $\tilde{\gamma}$ is a horizontal lift to S_k. To compute the curvature of this connection is easiest by introduction of the corresponding one-form, $\mathcal{A}_k$, on Σ_k.

For any given k, let us define a special vector field on $S_k \times U(1)$:

$$\chi_k|_{(\Psi,u)} := \frac{\mathrm{d}}{\mathrm{d}t}\left(\Psi \cdot \mathrm{e}^{-it},\ e^{ikt}u\right)\Big|_{t=0} = \mathcal{J} \oplus \left(-k\frac{\partial}{\partial\theta}\right),$$

where θ is the angular coordinate on $U(1)$; $u(\theta) = \mathrm{e}^{-\mathrm{i}\theta}$. Note that the bundle manifold Σ_k consists of integral curves of χ_k; a one-form on $S_k \times U(1)$ which is orthogonal to, and covariantly constant in the direction of χ_k, then defines a one-form on Σ. In this way, we introduce the connection form: define

$$\alpha_k := \frac{1}{2\hbar}\mathcal{J} \oplus (\mathrm{d}\theta), \tag{4.25}$$

where, here, $\mathcal{J}$ is viewed as a one-form on S_k. First, since S_k is the sphere of squared radius $2\hbar k$ (with respect to G), $\mathcal{J}^a\mathcal{J}_a = 2\hbar k$; therefore, $i_{\chi_k}\alpha_k = \frac{1}{2\hbar}\mathcal{J}_a\mathcal{J}^a - k = 0$. Next,

$$\begin{aligned} \pounds_{\chi_k}\alpha_k &= \left(i_{\chi_k}\mathrm{d} + \mathrm{d}i_{\chi_k}\right)\alpha_k = i_{\chi_k}\left(-\frac{1}{\hbar}{\pi_k}^*\omega_k \oplus 0\right) \\ &= -\frac{1}{\hbar}i_{\mathcal{J}}{\pi_k}^*\omega_k = 0. \end{aligned} \tag{4.26}$$

Thus, α_k defines a one-form $\mathcal{A}_k$ on Σ_k. Along the way, in Eqs. (4.26), we found that $\mathrm{d}\alpha_k = -\frac{1}{\hbar}{\pi_k}^*\omega_k \oplus 0$. As a consequence, $\mathrm{d}\mathcal{A}_k = -\frac{1}{\hbar}\tilde{\pi}_k^*\omega_k$. We have constructed the desired connection form; $\Sigma_k \xrightarrow{\tilde{\pi}_k} \mathcal{P}$ is the k^{th} pre-quantum bundle.

Now we must construct the line bundle which is associated to $\Sigma_k \xrightarrow{\tilde{\pi}_k} \mathcal{P}$; call it $E_k :\xrightarrow{\rho_k} \mathcal{P}$. One can simply use the definition provided in the appendix; i.e., define an equivalence relation on $\Sigma_k \times \mathbb{C}$ via $([\Psi, u]_k,\, c) \sim_k \left([\Psi, u]_k \cdot g^{-1},\, g \cdot z\right)$, and divide $\Sigma_k \times \mathbb{C}$ by this equivalence relation. Things are getting rather complicated; fortunately, there is a simpler way to go about this, which also makes closer contact with the original Hilbert space. We define an equivalence relation on $S_k \times \mathbb{C}$:

$$(\Psi, c) \sim_k (\Psi \cdot g^{-1},\, g^k \cdot c) \quad \forall\, g \in U(1), \tag{4.27}$$

and put $E_k := S_k \times \mathbb{C}/\sim_k$. Let us denote the equivalence class containing (Ψ, c) by $[\Psi, c]_k$; although this notation has been reserved for the principal bundle, the form of the equivalence relations are identical, so this should introduce no confusion. Naturally, the projection is defined in terms of π_k;

$$\rho_k : E_k \to \mathcal{P}, \qquad \rho_k : [\Psi, c]_k \mapsto \pi_k(\Psi).$$

Finally, given a local trivialization (U_i, ϕ_i) of $S_k \xrightarrow{\pi_k} \mathcal{P}$, one can introduce the obvious local trivializations of $E_k \xrightarrow{\rho_k} \mathcal{P}$;

$$\tau_i : U_i \times \mathbb{C} \to \rho_k^{-1}(U_i), \qquad \tau_i : (p, c) \mapsto [\phi(p, e),\, c]_k.$$

One then finds that the transition functions coincide with those for the principal bundle $\Sigma_k \xrightarrow{\tilde{\pi}_k} \mathcal{P}$, indicating that we have constructed the line bundle correctly.

We can now begin to turn the crank supplied by the geometric quantization procedure. Fix the integer k arbitrarily. Analogous to what was done in §4.1.1, we may associate, to every section $s : \mathcal{P} \to E_k$, a function on S_k:

$$s \circ \rho_k [\Psi, c]_k =: [\Psi, s^\star(\Psi)]_k. \tag{4.28}$$

It should be noted that this notation is not entirely consistent with that used in the general discussion of §4.1.1. The principal function was defined as a complex-valued function on the principal bundle to which the pre-quantum line bundle is associated; here $s^\star$ is a function on S_k—not Σ_k—but we will still call it the principal function associated to s. As a result of the definition, $[\Psi g, s^\star(\Psi g)]_k = [\Psi, g^k s^\star(\Psi g)]_k \Rightarrow$

$$s^\star(\Psi g) = g^{-k} s^\star(\Psi). \tag{4.29}$$

As a generalization of Eq. (4.3), to arbitrary integral k, we then have a one-to-one correspondence between sections of E_k and smooth functions on Σ_k which are, in the above sense, homogeneous of degree $-k$. The space of such functions is the pre-quantum state space,[3] $\mathcal{H}_{pre,\,k}$.

Let us now begin to consider the pre-quantum operators. We would like to describe the action of a pre-quantum operator $\hat{O}_{kf}$, on a section $s : \mathcal{P} \to E_k$, in terms of the induced action on the principal function $s^\star$; i.e., we desire an analogue of Eq. (4.8). Therefore, given any real function f

[3] Again, we proceed blindly. An inner-product will present itself at the end of the construction.

on $\mathcal{P}$, consider the fundamental vector field $f^{\sharp_k}$,

$$\begin{aligned} f^{\sharp_k}\big|_{[\Psi,c]_k} &:= \frac{\mathrm{d}}{\mathrm{d}t}[\Psi,c]_k \cdot \exp(-\mathrm{i}t f\circ\rho_k[\Psi,c]_k)\Big|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left[\Psi\exp\left(\frac{-\mathrm{i}t}{k}f\circ\pi_k(\Psi)\right),\ c\right]_k\Big|_{t=0}, \end{aligned} \tag{4.30}$$

on Σ_k. Since, again, we are attempting to make contact with the original Hilbert space, it should be useful to notice that $f^{\sharp_k}$ naturally defines the following vector field on S_k:

$$f^{\sharp_k'}\Big|_\Psi = \frac{\mathrm{d}}{\mathrm{d}t}\Psi\cdot\exp\left(\frac{-\mathrm{i}t}{k}f\circ\pi_k(\Psi)\right)\Big|_{t=0} = \frac{1}{k}\, f\circ\pi_k(\Psi)\mathcal{J}|_\Psi\,. \tag{4.31}$$

The useful fact is that similar to Eq. (4.7), multiplication by f, of a section s, corresponds to Lie differentiation of $s^\star$ with respect to $f^{\sharp_k'}$;

$$\begin{aligned} f^{\sharp_k'}(s^\star) &= \frac{1}{k}\,\frac{\mathrm{d}}{\mathrm{d}t}s^\star\left(\Psi\mathrm{e}^{-\mathrm{i}tf\circ\pi_k(\Psi)}\right)\Big|_{t=0} \\ &= \frac{1}{k}\,\frac{\mathrm{d}}{\mathrm{d}t}\left[\mathrm{e}^{\mathrm{i}ktf\circ\pi_k(\Psi)}s^\star(\Psi)\right]\Big|_{t=0} \\ &= \mathrm{i}f\circ\pi_k(\Psi)\, s^\star(\Psi). \end{aligned} \tag{4.32}$$

Now, let $\mathcal{D}_k$ denote the covariant derivative on Σ_k which is defined by the connection $\mathcal{A}_k$. As an immediate consequence of its definition,

$$(\mathcal{D}_{kX}s)^\star = \pounds_{\tilde{X}}s^\star,$$

where $\tilde{X}$ is the horizontal lift, to S_k, of X. In particular, let X_{kf} be the Hamiltonian vector field of f, as defined by the symplectic structure ω_k. Then the pre-quantum operator, $\hat{O}_{kf} : s \mapsto -\mathrm{i}\hbar\mathcal{D}_{kX_{kf}}s + fs$, acts on the principal function as

$$\left(\hat{O}_{kf}s\right)^\star = -\mathrm{i}\hbar\pounds_{\hat{X}_{kf}}s^\star, \qquad \hat{X}_{kf} = \tilde{X}_{kf} + \frac{1}{\hbar}f^{\sharp_k'}. \tag{4.33}$$

If $f : \mathcal{P} \to \mathbb{R}$ is an *observable* function (and hence a 'good' function), then by Eqs. (4.31) and (3.33), the vector field $\hat{X}_{kf}$, which is defined by operator $\hat{O}_{kf}$, is precisely the (restriction to S_k of the) Schrödinger vector field determined by the operator whose expectation value agrees with the function $f\circ\pi_k$!

Having endured all these mathematical constructions, we may now remember just a few facts: The pre-quantum states are represented by smooth functions $s^\star : S_k \to \mathbb{C}$ which are homogeneous of degree $-k$. To each function on $\mathcal{P}$ is associated a pre-quantum operator $\hat{O}_{kf}$, which acts on $s^\star$ (up to the factor of $-\mathrm{i}\hbar$) by differentiation with respect to the associated vector field $\hat{X}_{kf}$. It is an immediate consequence of the geometric quantization procedure that the mapping $X_{kf} \mapsto \hat{X}_{kf}$ preserves commutators of vector fields; this provides a formal homomorphism of the Poisson algebra of 'classical' functions into the commutator algebra of operators on $\mathcal{H}_{pre,k}$. Again, if f is observable, then the vector field $\hat{X}_{kf}$ is simply (the restriction to S_k of) the Schrödinger vector field of the operator $\hat{F}$ which defines f.

4.2.2 The Fock space

Having completed the pre-quantization procedure for general k, we simply generalize the procedure employed for $k = 1$. We begin by extending the pre-quantum principal functions to all of $\mathcal{H}$. For each $s^\star \in \mathcal{H}_{pre,\,k}$ we require that the extension $s^\star_{ext}$ satisfy

$$s^\star_{ext}(c\Psi) = \bar{c}^k s^\star(\Psi). \tag{4.34}$$

Any such extension is smooth on $\mathcal{H}$–$\{\mathbf{0}\}$, and the space of such functions on $\mathcal{H}$ is naturally isomorphic to $\mathcal{H}_{pre,\,k}$.

Just as in the previous case, the physical sections are those which are preserved by the Kähler polarization defined by (g_k, ω_k), where $g_{k\alpha\beta} = k g_{\alpha\beta}$, and similarly for ω_k. Of course, the polarization itself is independent of k; the polarization condition may then be written

$$\left(g_a{}^b - \mathrm{i} w_a{}^b\right) D_b s^\star = 0.$$

As in the previous section, one may easily write the derivatives of the extended principal functions with respect to the generators of the rays; they differ from Eqs. (4.12) and (4.13) by a factor of k, which does not affect the final statement of the extended polarization condition:

$$\left(G_a{}^b - \mathrm{i}\Omega_a{}^b\right) D_b s^\star_{ext} = 0. \tag{4.35}$$

Again, the pre-quantum states which satisfy the polarization condition are in one-to-one correspondence with homogeneous functions on $\mathcal{H}$, which are preserved by its Kähler polarization. The generalization, to arbitrary (integer) values of k, of the quantum conditions (4.15) is then

$$\begin{aligned} &i_k) \qquad s^\star_{ext}(c\Psi) = \bar{c}^k\, s^\star_{ext}(\Psi), \\ &ii_k) \qquad \left(G_b{}^a - \mathrm{i}\Omega_b{}^a\right) D_a s^\star_{ext} = 0 \qquad (\text{on } \mathcal{H}^\times). \end{aligned} \tag{4.36}$$

Physical states of the quantized k^{th} fake classical theory are functions on $\mathcal{H}$ which satisfy these two conditions.

We now argue that the general solution to conditions (i_k) and (ii_k) is provided by an element of the symmetric tensor product $\bigotimes^k_s \mathcal{H}$. Recall that the tensor product $\bigotimes^k \mathcal{H}$ consists of all bounded anti-linear maps $T : \mathcal{H} \times {}^{(\mathrm{k\ times})}\cdots \times \mathcal{H} \to \mathbb{C}$, and that $\bigotimes^k_s \mathcal{H}$ is the subset of $\bigotimes^k \mathcal{H}$ containing the totally symmetric elements.

Let us first establish some notation. For the moment, we will use an index notation for Hermitian tensor operations on $\mathcal{H}$. Here, Ψ^A and $\overline{\Psi}_A$ denote a (complex) vector in $\mathcal{H}$, and its dual[4] *with respect to the Hermitian inner-product*, respectively. The inner-product, $\langle \Phi, \Psi \rangle$, for example, is denoted $\overline{\Phi}_A \Psi^A$. Tensor operations are denoted similarly; for example, an element T of $\bigotimes^k \mathcal{H}$ may be written as $T^{A_1 \cdots A_k}$.

[4] One may prefer to think of $\overline{\Psi}_A$ as an element of the complex conjugate space $\overline{\mathcal{H}}$. Similarly, elements of $\bigotimes^k \mathcal{H}$ are actually tensors on $\overline{\mathcal{H}} \times {}^{(\mathrm{k\ times})}\cdots \times \overline{\mathcal{H}}$. The Riesz lemma allows us be sloppy in this regard. For a discussion of the abstract index notation appropriate for complex Hilbert spaces, see Wald [44].

Now let ζ be any complex-valued function on $\mathcal{H}$ which satisfies the conditions (i_k) and (ii_k). Let V be any finite-dimensional complex subspace of $\mathcal{H}$ and denote, by $\zeta|_V$, the restriction of ζ to V. Using exactly the same reasoning as given in §4.1.2, one can see that $\zeta|_V$ is an anti-holomorphic function on all of V, which is homogeneous of degree $-k$ in the sense of condition (i_k). There is a one-to-one correspondence between such functions and totally symmetric, rank k, tensors on V (or, more precisely, on $\overline{V}$). To see this, associate, to $\zeta|_V$, the tensor

$$T^{A_1\cdots A_k}_{\zeta|_V} := \frac{1}{k!}\frac{\partial}{\partial\overline{\eta}_{A_1}}\cdots\frac{\partial}{\partial\overline{\eta}_{A_1}}\zeta|_V,$$

where η denotes the generic element of V. $T^{A_1\cdots A_k}_{\zeta|_V}$ is clearly symmetric, and

$$\zeta|_V(\eta) = T^{A_1\cdots A_k}_{\zeta|_V}\overline{\eta}_{A_1}\cdots\overline{\eta}_{A_k}. \tag{4.37}$$

Conversely, given any such tensor, one may define, using Eq. (4.37), the corresponding function on V.

Therefore, given any solution ζ to equations (i_k) and (ii_k), one may associate, to each finite-dimensional complex subspace V of $\mathcal{H}$, a symmetric tensor $T^{A_1\cdots A_k}_{\zeta|_V}$. This uniquely defines a symmetric tensor $T^{A_1\cdots A_k}_{\zeta}$ on *all* of $\overline{\mathcal{H}}$; given elements $\overline{\Psi}_1,\cdots,\overline{\Psi}_k\in\overline{\mathcal{H}}$, one can simply construct the vector space V generated by the Ψ_i and define

$$T^{A_1\cdots A_k}_{\zeta}\overline{\Psi}_{1A_1}\cdots\overline{\Psi}_{kA_k} := T^{A_1\cdots A_k}_{\zeta|_V}\overline{\Psi}_{1A_1}\cdots\overline{\Psi}_{kA_k}.$$

Conversely, of course, any such tensor defines a solution to Eqs. (4.36).

In this manner, the k^{th} quantum Hilbert space—call it $\mathcal{H}_k$—is somewhat naturally identified with the symmetric tensor product $\bigotimes^k_s\mathcal{H}$. Recall that the symmetric Fock space is defined as

$$\mathcal{F}_s(\mathcal{H}) := \bigoplus_{k=0}^{\infty}\left(\bigotimes^k_s\mathcal{H}\right).$$

Therefore, $\mathcal{H}_k$ may be viewed as the 'k-particle' subspace; in particular, the k-particle states, sometimes written as $|\Phi_1,\cdots,\Phi_k\rangle$, each constituent 'particle' of which itself has a definite state in $\mathcal{H}$, are represented by

$$T^{A_1\cdots A_k} = \Phi^{A_1}_{(1}\cdots\Phi^{A_k}_{k)},$$

where $(1\cdots k)$ denotes symmetrization over all k labels. The standard description of second quantization is then recovered by quantization of *all* allowable symplectic structures on $\mathcal{P}$ at once!

To strengthen the sense in which the above statement holds, let us also verify that the second quantized observables agree with those provided by the standard Fock space quantization procedure. Let $\hat{F}$ be any observable quantum operator, and let f be the corresponding observable function, as defined by projection from the sphere S_k (i.e., $\pi_k^* f = \langle\hat{F}\rangle\big|_{S_k}$). Then, if $T^{A_1\cdots A_k}$ is the tensor

corresponding to the k-particle state ζ,

$$\begin{aligned}\left(\hat{O}_{kf}\zeta\right)(\Psi) &= -\mathrm{i}\hbar\, \pounds_{\hat{X}_{kf}}\zeta\Big|_{\Psi} \\ &= -\mathrm{i}\hbar \lim_{\varepsilon\to 0}\, T^{A_1\cdots A_k}\left(\overline{\Psi}_{A_1} + \frac{\mathrm{i}\varepsilon}{\hbar}\,\overline{\hat{F}\Psi}_{A_1}\right)\cdots\left(\overline{\Psi}_{A_k} + \frac{\mathrm{i}\varepsilon}{\hbar}\,\overline{\hat{F}\Psi}_{A_k}\right) \\ &= \sum_{i=1}^{k} T^{A_1\cdots A_k}\overline{\Psi}_{A_1}\,\overline{\Psi}_{A_1}\cdots\overline{\Psi}_{A_{i-1}}\,\overline{\hat{F}\Psi}_{A_i}\,\overline{\Psi}_{A_{i+1}}\cdots\overline{\Psi}_{A_k}.\end{aligned} \tag{4.38}$$

This should be recognized as the familiar form of the action of the second quantized operator on the k-particle subspace; its action on a direct product state may be somewhat heuristically written as

$$\hat{O}_{kf}\,|\Phi_1,\cdots,\Phi_k\rangle = \sum_{i=1}^{k}\left|\Phi_1,\cdots,\hat{F}\Phi_i,\cdots,\Phi_k\right\rangle. \tag{4.39}$$

5

Further Considerations

Let us now consider a few of the questions that provided the motivation for the geometric formulation of quantum mechanics.

The structures involved in the geometric description of quantum mechanics are much more easily generalized than those of the algebraic description. There are three elements of standard quantum mechanics which immediately come to mind: the phase space, the set of observables, and the dynamics. While, there is no clear idea how to construct a "non-linear Hilbert space", one might entertain the idea of a generalized quantum theory whose phase space is an arbitrary Kähler manifold, $(\mathcal{M}, g, \omega)$. In the standard picture, the observables preserve the metric; this is a strong constraint, which is satisfied only on a 'small' subset of the class of all smooth functions on the phase space. One may wish to consider the set of *all* functions on the generalized phase space $\mathcal{M}$, to represent the 'observable' quantities. Lastly, while the dynamics of standard quantum mechanics preserves both the metric and the symplectic structure, one might consider a generalized dynamics which preserves only the symplectic structure. For example, the flow associated to an arbitrary smooth function on $\mathcal{M}$ will be of this nature. These three features of the formalism may be extended independently of one another. The extension of the notion of the spectrum and its physical interpretation are not considered in this thesis; nor do we examine any other aspects of the probabilistic interpretation of such a radically generalized quantum mechanics. It should be pointed out that a complete and consistent formulation of a generalized mechanics which includes a suitable measurement theory will not be possible in general. Quantum mechanics is very rigid in this regard, and any generalization will have to be *finely* tuned to incorporate the essential features of the standard picture.

It is partly for this reason that extending a physical formalism in a *reasonable* way can be very difficult. Generalizations are easy to find, but they are typically of no physical relevance. Therefore, it is essential that one has a firm grasp of the character of the original theory. The next section is devoted to the study of what picks out quantum mechanics from all generalized formalisms. We will find that the essential feature of ordinary quantum mechanics is the maximality (in a sense to be explained) of the set of observables.

Having come to an understanding of what sets the standard formalism apart from the natural

generalized frameworks, we take a brief look at some of the familiar considerations that have been made in the past. A number of researchers have considered non-linear modifications of the quantum dynamics which do not alter the kinematic structure; this simplest possible route to a generalized mechanics will be the primary focus of §5.2. We are not interested in suggesting a particular equation of motion; our interest is in a general framework in which to study such questions. Recently, Weinberg has also considered the question of such a general setting for non-linear modifications of quantum mechanics [6]. Unfortunately, since his formalism seems not to include many of the familiar examples, it does not have the flavor of a universal framework for the discussion of non-linear dynamics. We will see, however, that when viewed from a geometric perspective, there is a very strong sense in which Weinberg's formalism is quite universal. This is illustrated with two examples, which Weinberg had eliminated from consideration: the so-called "non-linear Schrödinger equation" and the logarithmic equation of Bialynicki-Birula and Mycielski.

The last section is devoted to some preliminary questions involving the relationship between classical and quantum mechanics. Since quantum mechanics admits a description in terms of symplectic geometry, it is natural to ask for the relationship between the symplectic structures of classical and quantum mechanics. This question is considered for the particular case in which the corresponding classical phase space is linear. We find that the quantum phase space may be viewed as a bundle over the classical phase space, and notice that there is a natural decomposition of the tangent spaces of $\mathcal{P}$ into horizontal and vertical components. We will see that the classical symplectic structure may be obtained as the 'horizontal part' of the quantum one. The issue of dynamics is not a simple one; in order to gain some insights, we consider the harmonic oscillator. We are naturally led to a preferred embedding of the classical phase space into $\mathcal{P}$. After closer scrutiny, we find that this embedding is one of many horizontal sections of our quantum bundle. These sections, defined in purely geometric terms, are then seen to correspond precisely to generalized coherent state spaces. We then conclude with a brief look at semi-classical dynamics in terms of the WKB approximation.

5.1 Characterization of Ordinary Quantum Mechanics

In an attempt to generalize any mathematical formalism, it is important (if not essential) to have an understanding of what characterizes the original one. This gives one an idea of how dramatic any particular generalization may be and can provide a clearer indication of the physical role played by the relevant mathematical structures. This section presents a few results which characterize the kinematics of ordinary quantum mechanics.

Let us begin with an *arbitrary* connected Kähler manifold $(\mathcal{M}, g, \omega)$, which is to represent a radically generalized quantum theory, and ask for the conditions which guarantee that ordinary quantum mechanics is recovered. As we know, the observables of ordinary quantum mechanics are smooth functions on the phase space whose Hamiltonian vector fields are Killing; we therefore let the observables of the generalized framework consist of all smooth functions on $\mathcal{M}$ which preserve

the metric. Denote this set as

$$\mathcal{O} := \left\{ f : \mathcal{M} \to \mathbb{R} \;\middle|\; \pounds_{X_f} g = 0 \right\}.$$

Note that while we call elements of $\mathcal{O}$ *observables*, we have not introduced a complete generalized quantum formalism (and will not do so since it appears to be impossible at this level of generality).

It should be emphasized that there are Kähler manifolds which do not admit even a single observable function (other than the constants); the torus is an example. Let us briefly consider this fact. The discussion on page 39, which introduced the notion of symmetry data, is valid for *any* Kähler manifold. As before, we will denote by $\mathcal{S}_p$ the set of symmetry data at the point $p \in \mathcal{M}$ (if λ, X and K are zero-, one- and two-forms at p for which $\omega_\alpha{}^\gamma K_{\gamma\beta}$ is symmetric, then $(\lambda, X, K) \in \mathcal{S}_p$). Given an arbitrary $p \in \mathcal{M}$, any observable function $f : \mathcal{M} \to \mathbb{R}$ then determines an element of $\mathcal{S}_p$ (and is completely determined by its value and first two derivatives at the arbitrary point p). By Theorem 3.4 of §3.2.1, the converse also holds for ordinary quantum mechanics. In this sense, the phase space of standard quantum mechanics admits "as many observables as possible". This useful idea will be used extensively; let us introduce some terminology:

Definition 5.1. If, for each $p \in \mathcal{M}$, every element $(\lambda, X, K) \in \mathcal{S}_p$ is integrable in the sense of Theorem 3.4, we will say that $\mathcal{O}$ is *maximal*.

The set of observables of ordinary quantum mechanics is maximal; we will see that it is essentially the requirement of maximality that allows us to recover the standard formalism.

Before considering the implications of the maximality of $\mathcal{O}$, we recall some basic facts about the Riemann tensor of a Kähler manifold. Denote by $R_{\alpha\beta\gamma}{}^\delta$ the Riemann tensor of g; with our conventions, R is defined by Eq. (3.38). In addition to the Bianchi identity, the Riemann tensor on *any* differentiable manifold satisfies a few algebraic identities (e.g. $R_{[\alpha\beta\gamma]}{}^\delta = 0$, $R_{[\alpha\beta]\gamma}{}^\delta = R_{\alpha\beta\gamma}{}^\delta$, and $R_{\alpha\beta[\gamma\delta]} = R_{\alpha\beta\gamma\delta}$). By virtue of the Kähler structure of $\mathcal{M}$, the Riemann tensor possesses a few additional algebraic properties. First, since the complex structure $j^\beta{}_\alpha = \omega_\alpha{}^\beta$ is covariantly constant, for any k_δ we must have

$$R_{\alpha\beta\gamma}{}^\mu \omega_\mu{}^\delta k_\delta = 2\nabla_{[\alpha}\nabla_{\beta]}\omega_\gamma{}^\delta k_\delta = \omega_\gamma{}^\mu R_{\alpha\beta\mu}{}^\delta k_\delta.$$

Therefore,

$$R_{\alpha\beta\gamma\mu}\omega_\delta{}^\mu = R_{\alpha\beta\delta\mu}\omega_\gamma{}^\mu. \tag{5.1}$$

Contracting this equation with $\omega_\nu{}^\gamma$ gives another useful identity:

$$R_{\alpha\beta\mu\nu}\omega_\gamma{}^\mu\omega_\delta{}^\nu = R_{\alpha\beta\gamma\delta}. \tag{5.2}$$

The concept of scalar curvature is useful in the study of differential geometry of real manifolds, but not quite so useful for Kähler manifolds; any Kähler manifold (of complex dimension > 1) of constant curvature is necessarily flat [39]. There is, however, a generalization which we will find

extremely useful. The Riemann tensor is said to be of *constant holomorphic sectional curvature* (CHSC) at p if it is of the form

$$R_{\alpha\beta\gamma\delta}\big|_p = \frac{C}{2}\left(g_{\gamma[\alpha}g_{\beta]\delta} + \omega_{\alpha\beta}\omega_{\gamma\delta} - \omega_{\gamma[\alpha}\omega_{\beta]\delta}\right)\big|_p.$$

The real number C is the value of the holomorphic sectional curvature at p. If, for some constant C, the Riemann tensor assumes the above form at each point, $\mathcal{M}$ is called a manifold of CHSC$=C$. Recall from Eq. (3.39), that the projective Hilbert space is, with our conventions, a manifold of CHSC$=\frac{2}{\hbar}$. One can show that if the Riemann tensor on a *connected* Kähler manifold $\mathcal{M}$ is of CHSC at each point then it must be of overall CHSC [39]. We will always assume that $\mathcal{M}$ is connected.

Now that the preliminaries are out of the way, let us consider the observables of the general framework. The algebraic structure on the set of observables seems to be of great importance in both classical and quantum mechanics. In particular, the Lie structure is relevant for the description of the dynamics and for the consideration of integrable systems; the Jordan product of quantum observables is related to the notion of uncertainty, etc. The algebraic structure of $\mathcal{O}$ will be our primary concern in this section. Since the commutator of two Killing vector fields is another Killing field, the set of observables on an arbitrary Kähler manifold is closed under the Poisson bracket $\{,\}$. The Poisson bracket also satisfies the Jacobi identity; this equips $\mathcal{O}$ with the structure of a Lie algebra. On the entire function space, we may also define the commutative operation

$$\{f,k\}_+ := (f,k) + fk, \qquad (f,k) := \frac{\hbar}{2}(\nabla_\alpha f)g^{\alpha\beta}(\nabla_\beta k) = \frac{\hbar}{2}g(X_f, X_k).$$

Let us call $\{,\}_+$ the *symmetric bracket*. It is important to note that the set of observables on an arbitrary Kähler manifold is not necessarily closed under the symmetric bracket.

In order to thoroughly explore the algebraic structure of $\mathcal{O}$, we generalize a technique which was introduced by Ashtekar [45] for the study of isometries. Suppose $\underset{1}{f}$ and $\underset{2}{f}$ are any two observable functions on $\mathcal{M}$, and let $\underset{3}{f}$ be the Poisson bracket, $\underset{3}{f} = \{\underset{1}{f}, \underset{2}{f}\}$. Since the $\underset{i}{f}$ generate Killing vector fields, they determine elements $\left(\underset{i}{f}, \underset{i}{X}, \underset{i}{K}\right) \in \mathcal{S}_p$ for any $p \in \mathcal{M}$, where $\underset{i}{X} = X_{\underset{i}{f}}\big|_p$ and $\underset{i}{K}_{\alpha\beta} = \nabla_\alpha\left(X_{\underset{i}{f}}\right)_\beta|_p$. Of course, $\underset{3}{f} = \omega\left(\underset{1}{X}, \underset{2}{X}\right)$ and $\underset{3}{X}_\alpha = -g_{\alpha\beta}\left[X_{\underset{1}{f}}, X_{\underset{2}{f}}\right]^\beta\big|_p = \underset{2}{X}^\beta \underset{1}{K}_{\beta\alpha} - \underset{1}{X}^\beta \underset{2}{K}_{\beta\alpha}$, where $[,]$ denotes the Lie bracket of vector fields. Further, it is also easy to show that $\underset{3}{K}_{\alpha\beta} = \underset{2}{K}_\alpha{}^\gamma \underset{1}{K}_{\gamma\beta} - \underset{1}{K}_\alpha{}^\gamma \underset{2}{K}_{\gamma\beta} + \underset{1}{X}_\mu \underset{2}{X}_\nu R_{\alpha\beta}{}^{\mu\nu}$. This fact suggests that we define the following operation on the set $\mathcal{S}_p$ of symmetry data:

$$\begin{aligned}\left[\left(\underset{1}{f}, \underset{1}{X}, \underset{1}{K}\right),\ \left(\underset{2}{f}, \underset{2}{X}, \underset{2}{K}\right)\right]_p := \Big(&\omega\left(\underset{1}{X}, \underset{2}{X}\right),\ \ \underset{2}{X}^\beta \underset{1}{K}_{\beta\alpha} - \underset{1}{X}^\beta \underset{2}{K}_{\beta\alpha},\\ &\underset{2}{K}_\alpha{}^\gamma \underset{1}{K}_{\gamma\beta} - \underset{1}{K}_\alpha{}^\gamma \underset{2}{K}_{\gamma\beta} + \underset{1}{X}_\mu \underset{2}{X}_\nu R_{\alpha\beta}{}^{\mu\nu}\Big).\end{aligned} \tag{5.3}$$

This bracket was simply defined to produce the symmetry data corresponding to the Poisson bracket of $\underset{1}{f}$ and $\underset{2}{f}$. For the arbitrary Kähler manifold $\mathcal{M}$, it is straight-forward to check that $\mathcal{S}_p$ is closed under this bracket operation; however, the Jacobi identity will, in general, fail.

Before studying the algebraic structure defined by the above bracket on $\mathcal{S}_p$, let us make the analogous construction for the symmetric bracket; let $\underset{1}{f}, \underset{2}{f} \in \mathcal{O}$, and denote the corresponding elements of $\mathcal{S}_p$ by $(\underset{i}{f}, \underset{i}{X}, \underset{i}{K})$. Let k be the symmetric bracket, $k := \{\underset{1}{f}, \underset{2}{f}\}_+ = \underset{1}{f}\underset{2}{f} + \frac{\hbar}{2}\, g\big(\underset{1}{X}, \underset{2}{X}\big)$. Then one can easily show the following:

$$\begin{aligned}
\omega_\alpha{}^\beta \nabla_\beta k &= \underset{1}{f}\underset{2}{X}_\alpha + \underset{2}{f}\underset{1}{X}_\alpha + \frac{\hbar}{2}\omega_\alpha{}^\beta\big(\underset{1}{K}_{\beta\gamma}\underset{2}{X}^\gamma + \underset{2}{K}_{\beta\gamma}\underset{1}{X}^\gamma\big), \quad \text{and} \\
\nabla_\alpha\big(X_k\big)_\beta &= \underset{1}{f}\underset{2}{K}_{\alpha\beta} + \underset{2}{f}\underset{1}{K}_{\alpha\beta} + \hbar\, \underset{1}{K}_{\gamma[\alpha}\omega^{\gamma\delta}\underset{2}{K}_{\beta]\delta} + \underset{1}{X}^\mu\underset{2}{X}^\nu\left[\hbar R_{\alpha(\mu\nu)\gamma} - 2g_{\alpha(\mu}g_{\nu)\gamma}\right]\omega_\beta{}^\gamma.
\end{aligned}$$

Therefore it is natural to define the following commutative bracket:

$$\begin{aligned}
\Big(\big(\underset{1}{f}, \underset{1}{X}, \underset{1}{K}\big),\ \big(\underset{2}{f}, \underset{2}{X}, \underset{2}{K}\big)\Big)_p := \Big(&\underset{1}{f}\underset{2}{f} + \frac{\hbar}{2}\, g\big(\underset{1}{X}, \underset{2}{X}\big), \\
&\underset{1}{f}\underset{2}{X}_\alpha + \underset{2}{f}\underset{1}{X}_\alpha + \frac{\hbar}{2}\omega_\alpha{}^\beta\big(\underset{1}{K}_{\beta\gamma}\underset{2}{X}^\gamma + \underset{2}{K}_{\beta\gamma}\underset{1}{X}^\gamma\big), \\
&\underset{1}{f}\underset{2}{K}_{\alpha\beta} + \underset{2}{f}\underset{1}{K}_{\alpha\beta} + \hbar\, \underset{1}{K}_{\gamma[\alpha}\omega^{\gamma\delta}\underset{2}{K}_{\beta]\delta} + \underset{1}{X}^\mu\underset{2}{X}^\nu\left[\hbar R_{\alpha(\mu\nu)\gamma} - 2g_{\alpha(\mu}g_{\nu)\gamma}\right]\omega_\beta{}^\gamma\Big).
\end{aligned} \tag{5.4}$$

If $\mathcal{M}$ is a projective Hilbert space, this operation produces the symmetry data determined by the Jordan product of $\underset{1}{f}$ and $\underset{2}{f}$. For the generic case, $\mathcal{O}$ will not be closed under the symmetric bracket, and hence the bracket defined by Eq. (5.4) will not generally preserve the set of symmetry data (even if one assumes that $\mathcal{O}$ is maximal). Therefore, $(,)_p$ should be viewed as an operation on the space of triples (f, X, K), without the restriction imposed by Definition 3.2 on the two-form K. The condition that $\mathcal{S}_p$ is preserved under the action of $(,)_p$ is quite a strong one:

Lemma 5.1. *The set, $\mathcal{S}_p$, of symmetry data is closed under $(,)_p$ if and only if the Riemann tensor is of CHSC=$\frac{2}{\hbar}$ at p.*

Proof. If $R\big|_p$ is of CHSC=$\frac{2}{\hbar}$, then it is rather simple to verify the closure of $\mathcal{S}_p$. We consider only the converse. Let $(\underset{i}{f}, \underset{i}{X}, \underset{i}{K}) \in \mathcal{S}_p$ be arbitrary ($i = 1, 2$), and put

$$L_{\alpha\beta} := \underset{1}{f}\underset{2}{K}_{\alpha\beta} + \underset{2}{f}\underset{1}{K}_{\alpha\beta} + \hbar\, \underset{1}{K}_{\gamma[\alpha}\omega^{\gamma\delta}\underset{2}{K}_{\beta]\delta} + \underset{1}{X}^\mu\underset{2}{X}^\nu\left[\hbar R_{\alpha(\mu\nu)\gamma} - 2g_{\alpha(\mu}g_{\nu)\gamma}\right]\omega_\beta{}^\gamma.$$

L is the second rank tensor at p which is defined by Eq. 5.4. The $(,)_p$ bracket then preserves $\mathcal{S}_p$ if and only if L is necessarily skew-symmetric and $\omega_\alpha{}^\gamma L_{\gamma\beta} = \omega_\beta{}^\gamma L_{\gamma\alpha}$. That the first three terms individually satisfy these conditions is easy to show; let us focus on the third term. Put

$$S_{\alpha\beta\mu\nu} := \left[\hbar R_{\alpha(\mu\nu)\delta} - 2g_{\alpha(\mu}g_{\nu)\delta}\right]\omega_\beta{}^\delta.$$

That $\omega_\gamma{}^\beta S_{\alpha\beta\mu\nu} = \omega_\alpha{}^\beta S_{\gamma\beta\mu\nu}$ quickly follows from the algebraic properties of R; therefore, the second condition is satisfied. We need only show that $S_{\alpha\beta\mu\nu} + S_{\beta\alpha\mu\nu} = 0$. Contraction of this equation with $\omega_\delta{}^\beta$ gives, after a short manipulation, the equation

$$0 = \left(g_\delta{}^\beta g_\alpha{}^\gamma - \omega_\delta{}^\beta\omega_\alpha{}^\gamma\right)\left(R_{\mu\beta\nu\gamma} + R_{\nu\beta\mu\gamma}\right) + \frac{4}{\hbar}\left(g_{\alpha(\mu}g_{\nu)\delta} + \omega_{\alpha(\mu}\omega_{\nu)\delta}\right).$$

If one now skew-symmetrizes with respect to α and μ and uses the algebraic properties of the Riemann tensor, one obtains the desired result:

$$R_{\alpha\beta\mu\nu} = \tfrac{1}{\hbar}\left(g_{\mu[\alpha}g_{\beta]\nu} + \omega_{\alpha\beta}\omega_{\mu\nu} - \omega_{\mu[\alpha}\omega_{\beta]\nu}\right). \qquad \square$$

Note that the above consideration involves only a single point of $\mathcal{M}$. If the set of observables is maximal, then its elements are in one-to-one correspondence with elements of $\mathcal{S}_p$ for any $p \in \mathcal{M}$ (by definition). Therefore, if $\mathcal{O}$ is maximal, its closure under the symmetric bracket is equivalent to closure of $\mathcal{S}_p$ under $(,)_p$, for an arbitrary $p \in \mathcal{M}$. If we apply Theorem 5.1 to each point of $\mathcal{M}$, we immediately obtain

Corollary 5.2. *Suppose $\mathcal{O}$ is maximal. Then $\mathcal{O}$ is closed under the symmetric bracket if and only if the Riemann tensor is of CHSC=$\frac{2}{\hbar}$.*

Before considering the skew-symmetric bracket on $\mathcal{S}_p$, let us make a preliminary observation. Recall that the triple (f, X, K) is an element of $\mathcal{S}_p$ iff K is a two-form at p for which $\omega_\alpha{}^\gamma K_{\gamma\beta} = \omega_\beta{}^\gamma K_{\gamma\alpha}$. This condition can be nicely stated as follows. Define

$$\mathbb{P}_{\pm\alpha\beta}{}^{\mu\nu} := \frac{1}{2}\left(g_{[\alpha}{}^{\mu}g_{\beta]}{}^{\nu} \pm \omega_{[\alpha}{}^{\mu}\omega_{\beta]}{}^{\nu}\right). \tag{5.5}$$

One can easily verify the following:

$$\mathbb{P}_{\pm\alpha\beta}{}^{\gamma\delta}\mathbb{P}_{\pm\gamma\delta}{}^{\mu\nu} = \mathbb{P}_{\pm\alpha\beta}{}^{\mu\nu},$$
$$\mathbb{P}_{\pm\alpha\beta}{}^{\mu\nu}\mathbb{P}_{\mp\alpha\beta}{}^{\mu\nu} = 0.$$

Therefore $\mathbb{P}_\pm$ are *projections* on the space of two-forms; if L is a two-form, then

$$\left(\mathbb{P}_+ L\right)_{\alpha\beta} = L_{\alpha\beta} \;\Leftrightarrow\; \left(\mathbb{P}_- L\right)_{\alpha\beta} = 0 \;\Leftrightarrow\; \omega_\alpha{}^\gamma L_{\gamma\beta} = \omega_\beta{}^\gamma L_{\gamma\alpha}.$$

Now let us consider the skew-symmetric bracket, $[,]_p$. Recall that while this bracket preserves $\mathcal{S}_p$, it is not generally a Lie bracket. The satisfaction of the Jacobi identity constrains the Riemann tensor at p nearly as strongly as does the closure of $\mathcal{S}_p$ under the commutative operation considered above:

Lemma 5.3. $[,]_p$ *is a Lie bracket on $\mathcal{S}_p$ if and only if the Riemann tensor is of CHSC at p.*

Proof. If R is of CHSC, then one can easily verify that the bracket satisfied the Jacobi identity. We will prove the converse by brute force. Let $\underset{i}{\Lambda} = \left(\underset{i}{f}, \underset{i}{X}, \underset{i}{K}\right)$ be arbitrary elements of $\mathcal{S}_p$, $(i = 1, 2, 3)$. Using the definition (5.3), one finds that

$$\left[\underset{1}{\Lambda}, \left[\underset{2}{\Lambda}, \underset{3}{\Lambda}\right]_p\right]_p + \left[\underset{2}{\Lambda}, \left[\underset{3}{\Lambda}, \underset{1}{\Lambda}\right]_p\right]_p + \left[\underset{3}{\Lambda}, \left[\underset{1}{\Lambda}, \underset{2}{\Lambda}\right]_p\right]_p = \left(0,\, 0,\, M_{\alpha\beta}\right),$$

where

$$M_{\alpha\beta} = \sum_{\text{cyclic perm's}} 2\underset{1}{X}{}^{\mu}\underset{2}{X}{}^{\nu}\underset{3}{K}{}_{\gamma\delta}\left[R_{\alpha\beta[\mu}{}^{\gamma}g_{\nu]}{}^{\delta} + R_{\mu\nu[\alpha}{}^{\gamma}g_{\beta]}{}^{\delta}\right],$$

and the sum is over cyclic permutations of the labels $1, 2, 3$. We would like to show that $M_{\alpha\beta}$ vanishes. Using the projection defined above, we may write M as

$$M_{\alpha\beta} = \sum_{\text{cyclic perm's}} 2\underset{1}{X}^{\mu}\underset{2}{X}^{\nu}\underset{3}{T}_{\sigma\tau}\mathbb{P}_{+\gamma\delta}{}^{\sigma\tau}\left[R_{\alpha\beta[\mu}{}^{\gamma}g_{\nu]}{}^{\delta} + R_{\mu\nu[\alpha}{}^{\gamma}g_{\beta]}{}^{\delta}\right].$$

Notice that we may view the $\underset{i}{T}$ as *arbitrary* tensors at p; therefore M is guaranteed to vanish iff

$$0 = \mathbb{P}_{+\gamma\delta}{}^{\sigma\tau}\left[R_{\alpha\beta[\mu}{}^{\gamma}g_{\nu]}{}^{\delta} + R_{\mu\nu[\alpha}{}^{\gamma}g_{\beta]}{}^{\delta}\right].$$

For the moment, assume that $\mathcal{M}$ is of finite complex dimension n; this assumption will be eliminated shortly. We may then contract the above equation with the tensor $g_{\tau}{}^{\nu}\omega_{\delta\sigma}\omega_{\gamma}{}^{\mu}$, which yields, after plenty of inspired manipulation,

$$(*) \qquad (n+1)R_{\alpha\beta\gamma\delta} = R_{\gamma[\alpha}g_{\beta]\delta} - \frac{1}{2}R_{\alpha\beta\mu}{}^{\nu}\omega_{\nu}{}^{\mu}\omega_{\gamma\delta} - \omega_{\gamma}{}^{\mu}R_{\mu[\alpha}\omega_{\beta]\delta},$$

where $R_{\alpha\beta} := R_{\alpha\gamma\beta}{}^{\gamma}$ is the Ricci tensor. Contraction of $(*)$ with $\omega^{\gamma\delta}$ yields $R_{\alpha\beta\gamma\delta}\omega^{\gamma\delta} = -2R_{\gamma[\alpha}\omega_{\beta]}{}^{\gamma}$, which, when substituted back into $(*)$, produces the expression

$$(**) \qquad (n+1)R_{\alpha\beta\gamma\delta} = R_{\gamma[\alpha}g_{\beta]\delta} - R_{\mu[\alpha}\omega_{\beta]\mu}\omega_{\gamma\delta} - \omega_{\gamma}{}^{\mu}R_{\mu[\alpha}\omega_{\beta]\delta}.$$

Finally, contraction of this equation with $g^{\alpha\gamma}$ yields the Ricci tensor, $R_{\alpha\beta} = (R^{\gamma}{}_{\gamma}/2n)g_{\alpha\beta}$, which, when substituted back into $(**)$ yields the desired result:

$$R_{\alpha\beta\gamma\delta} = \frac{R}{2n(n+1)}\left[g_{\gamma[\alpha}g_{\beta]\delta} + \omega_{\alpha\beta}\omega_{\gamma\delta} - \omega_{\gamma[\alpha}\omega_{\beta]\delta}\right].$$

This proves our claim for the finite-dimensional case. It is not difficult to extend the result to infinite-dimensions. Let V be any finite-dimensional complex subspace of $T_p\mathcal{M}$. Then we can restrict the data $\underset{i}{\Lambda}$ to V and apply the above argument to see that the restriction of the Riemann tensor to any finite-dimensional complex subspace is of the desired form. As a consequence, R itself is of CHSC at p. □

Again, if $\mathcal{O}$ is maximal then $[,]_p$ is a Lie bracket on $\mathcal{S}_p$ for any $p \in \mathcal{M}$. As a consequence of Lemma 5.3, the Riemann tensor is of CHSC at each point of $\mathcal{M}$, and therefore $\mathcal{M}$ is a manifold of CHSC; this is summarized by

Corollary 5.4. *Suppose $\mathcal{O}$ is maximal. Then $\mathcal{M}$ is a manifold of CHSC.*

Note that these results do *not* specify the (constant) value of the holomorphic sectional curvature, but merely states that the Riemann tensor assumes this special form. However, by use of Lemma 5.1, we have

Corollary 5.5. *Suppose that $\mathcal{O}$ is maximal and that $\mathcal{S}_p$ is closed under $(,)_p$ for a single point $p \in \mathcal{M}$. Then $\mathcal{M}$ is a manifold of CHSC$=\frac{2}{\hbar}$.*

Naturally, we would like to go a step further with the above corollary and obtain conditions which guarantee that $\mathcal{M}$ is a projective Hilbert space. Any two finite-dimensional Kähler manifolds which are complete, simply-connected and of CHSC$=\frac{2}{\hbar}$ are isomorphic [39, Ch. 4]. Therefore, in the finite-dimensional case, we have obtained a characterization of the structure that picks out the standard quantum kinematics from all possible generalized frameworks. If the generalized phase space is a complete, simply-connected Kähler manifold and the set of observables (consisting of all smooth functions which preserve the metric) is maximal and closed under $\{,\}_+$, then one is dealing with the structure of ordinary quantum mechanics. It is not clear that one may do the same sort of thing for the infinite-dimensional case. Indeed there may exist *many* different infinite-dimensional Kähler manifolds (satisfying the above completeness requirements) of the same constant holomorphic sectional curvature. If this is indeed the case, we would be presented with generalizations of the standard quantum kinematics which seem to be quite subtle. Such minor generalizations are likely to admit a consistent measurement theory and could therefore be quite interesting.

5.2 Generalized Dynamics

Let us now turn to some considerations which are a bit more familiar. In the past, a number of researchers have considered non-linear modifications of the Schrödinger equation, but have retained the kinematic structure. All of these "non-linear Schrödinger equations" have the feature that the norm of the state vector is preserved by the generalized evolution. We will see that another commonality (which is not at all obvious) is the fact that these familiar generalized dynamics' induce Hamiltonian flows on the projective space. We will be naturally led to a picture that has been suggested by Weinberg [6]. Moreover, we will find that Weinberg's generalization (or, strictly speaking, class of generalizations) is universal in a precise sense. Many of the familiar modifications of Schrödinger's equation do not seem to be of the Weinberg type. However, as is explained in §5.2.2, they are equivalent to dynamical equations which *are* compatible with Weinberg's formalism. The nature of Weinberg's framework will then be seen to be much more general than it initially appears. In fact, this generality of the formalism is probably the reason that it is physically unacceptable. As is discussed by Czachor [46], the Weinberg formalism can lead to faster-than-light signals. Our primary intent in this section is not to support the idea, but to explain its geometrical meaning and illustrate the clarity that is introduced by the geometric viewpoint.

5.2.1 Weinberg's generalized quantum mechanics

Let us begin by considering the space of all smooth functions on the projective space. Let us think of $\mathcal{P}$ as a reduced phase space, as described in §3.1.1. With this viewpoint, one considers the observables as functions on S (the unit sphere in $\mathcal{H}$, with respect to $\langle,\rangle$) which are constant along the integral curves of $\mathcal{J}$; let us denote this set by

$$\mathcal{O}_W := \{f : S \to \mathbb{R} \mid \pounds_{\mathcal{J}} f = 0\}. \tag{5.6}$$

In this section, by the term 'observable' we mean an element of $\mathcal{O}_W$. Recall that the flow on $\mathcal{P}$ which is induced by $f \in \mathcal{O}_W$ may be obtained as follows. First, extend f to a neighborhood of S *arbitrarily*. Next, construct the Hamiltonian vector field determined by this extended function and restrict to S. The horizontal part of this vector field is independent of the particular extension and projects to a vector field on $\mathcal{P}$ (c.f. Eq. (3.8)).

There is a fairly convenient (but, by no means necessary) convention by which we may choose to extend an element of $\mathcal{O}_W$ to all of $\mathcal{H}^\times$. Suppose that $\hat{F}$ is a bounded, self-adjoint operator on $\mathcal{H}$, and that f is the restriction of its expectation value to S. One may consider it fairly natural to agree to a convention with respect to which the extended f (call it f_{ext}) is simply $\langle \hat{F} \rangle$;

$$f_{ext}(\Psi) := ||\Psi||^2 f(\Psi/||\Psi||). \tag{5.7}$$

This equation may be used to extend *any* element of $\mathcal{O}_W$. Notice that the extension of any $f \in \mathcal{O}_W$, commutes *strongly* with the constraint of Eq. (3.4) and hence the flow along $X_{f_{ext}}$ preserves the norm of the state vector. Note also that for any $f \in \mathcal{O}_W$, the Hamiltonian vector field $X_{f_{ext}}$ is homogeneous of degree one,

$$X_{f_{ext}}(c\Psi) = cX_{f_{ext}}(\Psi) \quad \forall\, c \in \mathbb{C};$$

the flow on $\mathcal{H}^\times$ which is generated by f_{ext} is homogeneous, but not linear.

Let us take a brief digression. We have extended the set of observables quite dramatically. In the finite-dimensional case, for example, $\mathcal{O}_W$ is infinite-dimensional, while the standard set $\mathcal{O}_{qu}$ is finite-dimensional. It seems natural to wonder whether there is some natural extension of $\mathcal{O}_{qu}$ which is not so 'huge'. For example, one might desire that the extended set of observables generate flows which, though generally non-linear, preserve the Hermitian inner-product. Denote by Φ_t the motion along the integral curve of $X_{f_{ext}}$ which passes through Φ at $t = 0$ (i.e., $\frac{\mathrm{d}}{\mathrm{d}t}\Phi_t = X_{f_{ext}}\big|_{\Phi_t}$). Given two points $\Psi, \Phi \in \mathcal{H}^\times$, we ask whether $\langle \Psi_t, \Phi_t \rangle$ is constant. We see that

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \Psi_t,\, \Phi_t \rangle\Big|_{t=0} = \frac{1}{2\hbar}\left[G(X_{f_{ext}}\big|_\Psi,\, \Phi) + G(\Psi,\, X_{f_{ext}}\big|_\Phi)\right] + \frac{\mathrm{i}}{2\hbar}\left[\Omega(X_{f_{ext}}\big|_\Psi,\, \Phi) + G(\Psi,\, X_{f_{ext}}\big|_\Phi)\right].$$

This is to vanish iff $G(X_{f_{ext}}\big|_\Psi,\, \Phi) = -G(\Psi,\, X_{f_{ext}}\big|_\Phi)$. Since the left side is linear in Φ, so must be the right side. Thus, $\Phi \mapsto X_{f_{ext}}\big|_\Phi$ must be a skew-self-adjoint linear map. Though we haven't proven it, it seems intuitively clear that the result holds for the unbounded case as well. There is no such thing as a non-linear unitary flow on $\mathcal{H}$; we seem to be stuck with the set $\mathcal{O}_W$.

Having decided to extend all functions as in Eq. (5.7), we see that there is a one-to-one correspondence between elements of $\mathcal{O}_W$ and smooth functions on $\mathcal{H}^\times$ which are gauge-invariant and homogeneous of degree two (if $\mathcal{H}$ is viewed as a vector space over the complexes, this corresponds to homogeneity of degree one in Ψ and $\overline{\Psi}$). Motivated, as we are, by the desire for a general framework for non-linear generalizations of quantum mechanics, Weinberg considered precisely this set of homogeneous functions on $\mathcal{H}^\times$ as generalized observables. Let us then call the f_{ext} *Weinberg functions*. We have a one-to-one correspondence between Weinberg functions and smooth functions on the projective Hilbert space.

Next, the dynamical evolution considered in [6] corresponds to integration of the Hamiltonian vector field of one of these homogeneous functions (we will not worry about the important fact that the evolution will typically correspond to a vector field which is only densely defined). Weinberg's picture is looking very natural from the geometric point of view. However, there is one very important point to note. Weinberg's motivation for the homogeneity condition stems from the desire for Ψ and $c\Psi$ to represent the same physical state for any $c \in \mathbb{C}$. This way, if $\Psi(t)$ solves the equation of motion, then so does $c\Psi(t)$ for any $c \in \mathbb{C}$. Our attitude toward the homogeneity of the observables is one of mere convention; if $\Psi \in S$ and $|c| \neq 1$, then $c\Psi$ does not represent any physical state at all! Our viewpoint is one in which the only points of physical significance are those of S. This may seem a very minor point, but it has implications which provide a sense in which Weinberg's framework becomes universal.

5.2.2 Universality of the Weinberg formalism

To make closer contact with previous work, it should be useful to reinstate a complex notation. For the rest of this section, we will be interested in the non-relativistic mechanics of a point particle moving in $\mathbb{R}^n$. The Hilbert space is represented by the space of square-integrable functions on R^n. We will now view the generalized observables as functions on $\mathcal{H}^\times$, i.e., by the statement $f \in \mathcal{O}_W$ we will mean that $f : \mathcal{H}^\times \to \mathbb{R}$ is one of the preferred homogeneous functions discussed above. The Hamiltonian vector field corresponding to f corresponds to the functional

$$X_f[\Psi](x) = \frac{1}{\mathrm{i}\hbar}\frac{\delta f}{\delta \overline{\Psi}}(x).$$

Of the familiar non-linear modifications to the Schrödinger equation, the best known is that which goes by the name "the non-linear Schrödinger equation". This equation is given by

$$\mathrm{i}\hbar\frac{\partial \Psi}{\partial t}(x,t) = \big(\hat{H}_0\Psi\big)(x,t) + \varepsilon|\Psi(x,t)|^2\Psi(x,t), \tag{5.8}$$

where $\hat{H}_0 = \frac{1}{2m}\hat{P}^2 + \hat{V}$ is the standard Hamiltonian operator describing a non-relativistic particle under the influence of a conservative force. Note that Weinberg correctly pointed out that the "results obtained by the mathematical studies of this equation are unfortunately of no use to us here", due to the fact that the homogeneity requirement is not satisfied. Let us consider it a bit further anyway. First, by dividing this equation by $\mathrm{i}\hbar$ we obtain the vector field on $\mathcal{H}^\times$ which corresponds to the generalized flow:

$$X[\Psi](x) = X_{H_0}[\Psi](x) + X_\varepsilon[\Psi](x), \qquad X_\varepsilon[\Psi](x) = \frac{\varepsilon}{\mathrm{i}\hbar}|\Psi(x,t)|^2\Psi(x,t).$$

We are now beginning to be a bit sloppy; since $\hat{H}_0$ is a genuinely self-adjoint operator (it is not globally defined), the first term is not quite a Hamiltonian vector field. The notation is helpful, however, and we will continue with the minor abuse.

The first interesting observation is the fact that the non-linear term in the above expression is actually the Hamiltonian vector field determined by the function

$$h_\varepsilon(\Psi) := \frac{\varepsilon}{2}\int \mathrm{d}^n y\ \left[\overline{\Psi(y,t)}\Psi(y,t)\right]^2,$$

i.e., $\frac{1}{\mathrm{i}\hbar}(\delta h_\varepsilon/\delta\overline{\Psi})(x) = X_\varepsilon[\Psi](x)$. The implication of this fact is the following: being interested only with the flow on $\mathcal{P}$, we may may feel free to restrict h_ε to S. The restriction $h_\varepsilon\big|_S$ contains all of the information about h_ε needed for the consideration of the flow on $\mathcal{P}$. Of course, to construct a vector field on S which projects to the relevant one on $\mathcal{P}$, we extend $h_\varepsilon\big|_S$ in any desired manner. In particular, we may extend it in the way which Weinberg would suggest:

$$h_\varepsilon^{'}(\Psi) := \|\Psi\|^2\, h_\varepsilon\big|_S(\Psi/\|\Psi\|)$$

Therefore, the flow on $\mathcal{P}$ which is defined by h_ε may also be described by a Weinberg function!

The non-linear Schrödinger equation is a fairly trivial example since the generating function h_ε is itself homogeneous, but only of the wrong degree. Let us consider another example which is not so obvious.

The Bialynicki-Birula equation

In an effort to tackle the problem of combining systems which are subject to a non-linear equation of motion, Bialynicki-Birula and Mycielski [7] were naturally led to a *logarithmic* equation of motion. The classical configuration space of the entire system is $\mathbb{R}^n$ and the dynamics is assumed to be of the general form

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi(x,t) = (\hat{H}_0\Psi)(x,t) + F\big(\big|\Psi(x,t)\big|^2\big)\Psi(x,t), \tag{5.9}$$

where $\hat{H}_0 = -\frac{\hbar^2}{2m}\Delta + V(x,t)$. Here, Δ is the Laplace operator on $\mathbb{R}^n$ and V is, of course, real-valued. While they provide an interesting discussion of the character of the above dynamical equation (for example, the norm of the state-vector is preserved), we shall only be concerned with their main result: suppose that the potential energy is of the form $V(x,t) = V_1(x_1,\ldots,x_k,t) + V_2(x_{k+1},\ldots,x_n,t)$. Then there is only one function F which does not introduce an interaction between the two sub-systems defined by the form of V. The evolution of the state $\Psi(x,0) = \psi_1(x_1,\ldots,x_k,0)\psi_2(x_{k+1},\ldots,x_n,0)$ may be obtained by evoloving ψ_1 and ψ_2 separately (according to the lower-dimensional analogues of Eq. (5.9)) if and only if

$$F(\rho) = -b\ln(a^n\rho),$$

for some constants a and b. Note that a is a length scale, but has no physical significance since it may be altered by addition of a constant to the Hamiltonian operator.

Let us then take F to be of the Bialynicki-Birula, Mycielski form. Choosing units with respect to which $a = 1$ and assuming V to be time-independent, the vector field on $\mathcal{H}^\times$ along which the system evolves is then given by

$$X[\Psi](x) = X_{H_0}[\Psi](x) + X_1[\Psi](x), \qquad X_1[\Psi](x) := -\frac{b}{\mathrm{i}\hbar}\ln\big(\big|\Psi(x)\big|^2\big)\Psi(x).$$

Once again, the non-linear term can easily be seen to be Hamiltonian;

$$X_1[\Psi] = X_{h_1}[\Psi], \qquad h_1(\Psi) := b\int \mathrm{d}^n y\,\overline{\Psi(y)}\Psi(y)\left[1 - \ln\left(\overline{\Psi(y)}\Psi(y)\right)\right].$$

We can then carry out the procedure used for the non-linear Schrödinger equation to see that the motion on the projective space is given by a Weinberg function. The resulting "homogeneous logarithmic equation" describing the flow on $\mathcal{H}^\times$ is determined by the Weinberg observable

$$\begin{aligned} h_1^{'}(\Psi) &= \|\Psi\|^2 h_1(\Psi/\|\Psi\|) \\ &= b\int \mathrm{d}^n y\, \overline{\Psi(y)}\Psi(y) - b\int \mathrm{d}^n y\, \overline{\Psi(y)}\Psi(y) \ln\left[\overline{\Psi(y)}\Psi(y)\right] \\ &\qquad + b\int \mathrm{d}^n y\, \overline{\Psi(y)}\Psi(y)\cdot \ln\Big[\int \mathrm{d}^n z\, \overline{\Psi(z)}\Psi(z)\Big] \\ &= h_1(\Psi) + b\|\Psi\|^2 \ln\left(\|\Psi\|^2\right). \end{aligned} \tag{5.10}$$

Even the logarithmic equation, which looks so dramatically 'inhomogeneous', induces a flow (or at least defines a vector field) on $\mathcal{P}$ which may be described by a function of Weinberg type. In retrospect, any generalized dynamics of the form specified by Eq. (5.9) may be written as a Hamiltonian flow (by 'integrating' the function F). One may then use the above procedure to write the induced flow on $\mathcal{P}$ in terms of a Weinberg function on $\mathcal{H}^\times$. This is true of any Hamiltonian flow on $\mathcal{H}^\times$ which leaves the unit sphere invariant.

By adopting the constrained system point of view, Weinberg's formalism is seen to be quite a general one. The author would be quite surprised to find a generalized dynamics with any interesting properties which may not be described, as above, in terms of Weinberg's language. An arbitrary Hamiltonian flow on $\mathcal{P}$ may be described in Weinberg's language and the Hamiltonian evolution seems the only reasonable kind for us to consider.

5.3 Semi-Classical Physics

We now consider another of our motivations for development of the geometric approach to quantum mechanics—the correspondence between classical and quantum mechanics. One might expect comparison of the classical and quantum kinematics to be easiest. This is the first topic of §5.3.1. For the simple case in which the classical phase space is linear, we come to an understanding of the relationship between the classical and quantum symplectic geometries. In fact, we will discover a number of natural embeddings of the classical phase space into the quantum one. This fact suggests an interesting approach to dynamical questions. In §5.3.1, we examine only the simplest case—the harmonic oscillator. Since the classical evolution is also linear in this case, we expect to find a fairly simple relationship between the classical and quantum dynamics. Of course, we do not expect to arrive at any fundamentally new conclusions; we simply hope to gain some insights regarding the role of geometry in the dynamical aspects of the classical limit. Along the way, we will have encountered a number of natural embeddings of the classical phase space into the quantum phase space. This fact suggests an interesting description of the coherent state spaces, which will be presented in §5.3.2. Finally, in §5.3.3 we briefly consider semi-classical dynamics from the standard point of view of Hamilton-Jacobi theory and the WKB method.

5.3.1 The classical limit

Many textbooks on quantum mechanics present two separate discussions of the so-called classical limit (see, e.g., [47]). One approach, is motivated by Hamilton-Jacobi theory and is related to the WKB approximation. The other is in terms of Ehrenfest's theorem, which describes the time dependence of the expectation value of an observable in terms of the Hamiltonian operator. The latter approach seems quite natural from our point of view, since in the geometric approach it is the expectation value functions that represent the observable quantities. Let us delay consideration of the dynamical questions and first consider the relationship between the classical and quantum kinematics.

Kinematics: Linear systems

We suppose that our quantum phase space $(\mathcal{P}, g, \omega)$ corresponds to a classical phase space (Γ, α) which is *linear*. We then have an association of the the elementary position and momentum variables of the classical theory to the basic position and momentum operators of the quantum theory. Let $\hat{Q}_i$ and $\hat{P}_i$ $(i=1, \ldots, n)$ be the basic quantum operators and denote the corresponding functions on $\mathcal{P}$ by q_i and p_i, respectively. Note that the Poisson brackets of these elementary quantities satisfy the same relations as do the classical variables:

$$\{q_i, p_j\} = \delta_{ij} \quad \text{and} \quad \{q_i, q_j\} = 0 = \{p_i, p_j\}.$$

Strictly speaking, however, we can not quite use these variables; since $\hat{Q}_i$, for example, is only densely defined, q_i is not a well-defined function on all of $\mathcal{P}$. The discussion below will therefore be heuristic, but the results will be justified in the end. Our non-rigorous presentation is desirable in that it provides an intuitive picture of the overall construction.

Let us denote the generic point of Γ by a pair (q_i^{cl}, p_i^{cl}), where q_i^{cl} and p_i^{cl} are, of course, the coordinates and momenta of the point which they label. The Ehrenfest approach to the classical limit suggests that to the point $x \in \mathcal{P}$, we identify the point $\big(q_i(x), p_i(x)\big) \in \Gamma$. This defines the mapping $\mathcal{P} \xrightarrow{\rho} \Gamma$ which takes a point of $x \in \mathcal{P}$ to the the element of Γ whose coordinates are given by the expectation values of the quantum operators at x. It will be convenient to let f_r $(r=1, \ldots, 2n)$ denote the generic elementary observable function on $\mathcal{P}$; i.e., $\{f_r\} = \{q_i, p_i, i=1, \ldots, n\}$. To describe the above projection another way, we may define an equivalence relation on $\mathcal{P}$: $x_1 \sim x_2 \Leftrightarrow f_r(x_1) = f_r(x_2)$ $\forall\, r=1, \ldots, 2n$. Then $\Gamma = \mathcal{P}/\sim$.

We have just described $\mathcal{P}$ as a bundle over the classical phase space Γ. This naturally defines a special subspace of of the tangent space $T_x\mathcal{P}$ which consists of tangent vectors which are vertical. That a vector $v \in T_x\mathcal{P}$ is vertical simply means that $v(f_r) = 0 \quad \forall\, r = 1, \ldots, 2n$; equivalently, $\omega\big(X_{f_r}|_x, v\big) = 0$ $\forall\, r=1, \ldots, 2n$. The vertical subspace may then be defined as

$$\mathcal{V}_x := \big\{ v \in T_x\mathcal{P} \;\big|\; \omega\big(X_{f_r}|_x, v\big) = 0 \;\; \forall\, r=1, \ldots, 2n \big\}. \tag{5.11}$$

For the moment, denote by $\big\langle X_{f_r} \big\rangle_x$ the linear span, at x, of the Hamiltonian vector fields of the basic quantities. The assignment of the vertical space to each point $x \in \mathcal{P}$ defines a vertical

distribution which is just the ω-orthogonal complement, $\mathcal{V}_x = \langle X_{f_r} \rangle_x^\perp$, of $\langle X_{f_r} \rangle_x$. Notice that since $\omega(X_{q_i}, X_{p_i}) = 1$, $\mathcal{V}_x$ does not intersect $\langle X_{f_r} \rangle_x$.

Let us show any tangent vector Y at x may be uniquely written as $Y = X + v$ for $v \in \mathcal{V}_x$ and $X \in \langle X_{f_r} \rangle_x$. Note that since $\omega(X_{q_i}, X_{p_i}) = 1$, and all other pairings vanish, $\{X_{f_r}|_x\}$ forms a canonical basis for $\langle X_{f_r} \rangle_x$. Given $Y \in T_x\mathcal{P}$, define $X := \sum_i [\omega(Y, X_{p_i}) X_{q_i} - \omega(Y, X_{q_i}) X_{p_i}]$ and put $v := Y - X$. Clearly $X \in \langle X_{f_r} \rangle_x$ and $Y = X + v$. To see that v is necessarily an element of $\mathcal{V}_x = \langle X_{f_r} \rangle_x^\perp$, notice that $\omega(X_{q_i}, v) = \omega(X_{q_i}, Y) - \omega(X_{q_i}, X) = \omega(X_{q_i}, Y) + \omega(Y, X_{q_i}) = 0$; similarly, $\omega(X_{p_i}, v) = 0$. Thus, $v \in \mathcal{V}_x$. That this decomposition of Y is unique should be obvious since X is uniquely determined by the symplectic product of Y with the X_{f_r}. As a consequence of this decomposition of the tangent spaces, on can easily find that $\mathcal{V}_x^\perp = \langle X_{f_r} \rangle_x^{\perp\perp} = \langle X_{f_r} \rangle_x$. All that one needs to obtain from this argument is the following; the tangent space $T_x\mathcal{P}$ has been decomposed as the sum $T_x\mathcal{P} = \mathcal{V}_x \oplus \mathcal{V}_x^\perp$ of two *symplectic* subspaces, where

$$\mathcal{V}_x^\perp = \{ X \in T_x\mathcal{P} \;\; | \;\; \omega(v, X) = 0 \;\; \forall \, v \in \mathcal{V}_x \}.$$

An element of $\mathcal{V}_x^\perp$ will be called a *horizontal* vector; horizontal vectors are simply those which are ω-orthogonal to the vertical directions. If Y is a vector *field* on $\mathcal{P}$ which is every horizontal (resp. vertical), we will simply write $Y \in \mathcal{V}^\perp$ (resp. $Y \in \mathcal{V}$). Note that if f is any algebraic function of the f_r, then $X_f \in \mathcal{V}^\perp$.

Let us make three more preliminary observations. First, the distribution $\mathcal{V}$ is clearly *integrable* since for any $v_1, v_2 \in \mathcal{V}$, $[v_1, v_2](f_r) = 0 \;\Rightarrow\; [v_1, v_2] \in \mathcal{V}$. Next, since $X_{f_r}(f_s) = \{f_s, f_r\}$, we must have $[v, X_{f_r}](f_s) = 0 \;\; \forall \, v \in \mathcal{V}$; so the X_{f_r} preserve vertical distribution. Similarly, the Hamiltonian vector field of any algebraic function of the f_r also preserves the distribution. Lastly, since $[X_{f_r}, X_{f_s}] = 0 \;\; \forall \, r, s$, the horizontal spaces are integrable; there exist horizontal sections of our quantum bundle over Γ.

Now we are prepared to reconstruct the classical symplectic structure. Let ξ and ζ be two vector fields on Γ, and let $\tilde{\xi}$ and $\tilde{\zeta}$ be their horizontal lift to $\mathcal{P}$, defined in the obvious fashion. Then the classical symplectic structure is given by

$$\alpha(\xi, \zeta) = \omega(\tilde{\xi}, \tilde{\zeta}); \tag{5.12}$$

alternatively, we may write $\rho^* \alpha = \omega_{\text{hor}}$. Given our particular construction in which Γ is a linear space, etc., it is obvious that this correctly defines the classical symplectic structure. However, in order to allow more general considerations in the future, let us see why this definition of α provides a symplectic structure on Γ. That α is well-defined is seen as follows: write $\tilde{\xi} = \xi^r X_{f_r}$ and $\tilde{\zeta} = \zeta^r X_{f_r}$. Then for any $v \in \mathcal{V}$,

$$\begin{aligned} \pounds_v [\omega(\tilde{\xi}, \tilde{\zeta})] &= \pounds_v [\omega(\xi^r X_{f_r}, \zeta^s X_{f_s})] \\ &= \pounds_v [\xi^r \zeta^s \{f_s, f_r\}] \;=\; 0, \end{aligned}$$

since $\{f_s, f_r\}$ is again one of the elementary functions and, for example, $\pounds_v \xi^r = 0$. Therefore, the definition is consistent. To see that α is non-degenerate, notice that if $\alpha(\xi, \zeta) = 0 \;\; \forall \, \zeta$, then

$\omega(\tilde{\xi}, \tilde{\zeta}) = 0 \ \ \forall\, \tilde{\zeta} \in \mathcal{V}^{\perp} \ \Rightarrow \ \tilde{\xi} \in \mathcal{V}$; by construction, $\tilde{\xi}$ must therefore vanish. Finally, that α is closed is clear since it may be obtained by pulling back ω via any horizontal section of $\mathcal{P} \xrightarrow{\rho} \Gamma$, and ω is closed.

The above arguments have been a bit dense; let us then summarize. For the case in which the quantum kinematics corresponds to a classical phase space which is linear, there is a fairly natural vertical distribution of subspaces of the tangent spaces of $\mathcal{P}$. The vectors in this distribution are simply those in directions along which the values of the basic functions (the q_i and p_i) are constant. The classical phase space is the space of leaves of this distribution; in this way, $\mathcal{P}$ is viewed as a bundle over the classical phase space. The symplectic structure on $\mathcal{P}$ defines a natural notion of horizontality, and the classical symplectic structure is simply the horizontal part of ω.

It would be desirable to extend the above construction to a more generic case. Our arguments do go a short step in this direction. Suppose that $\mathcal{V}$ is an integrable, symplectic distribution of $\mathcal{P}$ of (finite) co-dimension $2n$ which may be specified *locally* by the constancy of functions f_r, $r = 1, \ldots, 2n$. Suppose also that these functions may be chosen in such a way that their Poisson algebra closes (up to constants). Then the quotient of $\mathcal{P}$ by $\mathcal{V}$ inherits a natural symplectic structure.

Dynamics: The harmonic oscillator

Since the horizontal spaces introduced above are integrable, we may embed the classical phase space in $\mathcal{P}$ by choosing a horizontal section $s : \Gamma \to \mathcal{P}$. Since the vertical spaces are 'huge' there are an awful lot of these horizontal sections. It seems quite natural to ask whether any one of these is preferred in some sense. The answer to this question will certainly involve the dynamics, which we have not yet considered. In general, this is not likely to be trivial except for the simplest realistic case—the harmonic oscillator—which we will now consider.

Recall the situation encountered in the elementary treatment of the classical limit from the point of view of Ehrenfest's theorem. If $\hat{F}$ is any observable quantum operator, then the rate of change of its expectation value is given by $\frac{\mathrm{d}}{\mathrm{d}t}\langle \hat{F} \rangle = \frac{1}{i\hbar}\langle [\hat{F}, \hat{H}] \rangle$, where $\hat{H}$ is the Hamiltonian operator. Now let f and h be the corresponding observable functions on $\mathcal{P}$ (again, we ignore the fact that $\hat{H}$ is genuinely self-adjoint). Then, from the observations made in Ch. 3, we may write the above as $\frac{\mathrm{d}}{\mathrm{d}t} f = \{f, h\}$. The fact that this is exactly the classical expression is deceiving; as we know, there is not generally a manageable relationship between the Hamiltonian function $h : \mathcal{P} \to \mathbb{R}$ and the classical Hamiltonian function. In particular, h may not be written in terms of the q_i and p_i alone; this is true even for the harmonic oscillator. From the geometric point of view, this seems a bit puzzling, since the q_i and p_i 'follow the classical trajectories' if the Hamiltonian operator is quadratic. Let us understand this.

Our first task is to find the relationship between the quantum Hamiltonian and the classical

Hamiltonian function. This is not too difficult. The Hamiltonian operator is, of course,

$$\begin{aligned}\hat{H} &= \frac{1}{2m}\hat{P}^2 + \frac{m\omega^2}{2}\hat{Q}^2 \\ &= \frac{1}{2m}\{\hat{P},\,\hat{P}\}_+ + \frac{m\omega^2}{2}\{\hat{Q},\,\hat{Q}\}_+.\end{aligned} \tag{5.13}$$

Therefore, $h = \frac{1}{2m}p^2 + \frac{m\omega^2}{2}q^2 + \frac{1}{2m}(\Delta p)^2 + \frac{m\omega^2}{2}(\Delta q)^2$, where we have used Eq. (3.48). Terms such as the last two, involving the uncertainties, always arise when constructing the quantum Hamiltonian function; this is what provides the distinction between the classical and quantum dynamics in the generic case.

Notice that the Hamiltonian vector field X_h may be decomposed into two parts

$$X_h = X_{h_0} + X_{h_\Delta},$$

where h_0 is the 'classical Hamiltonian' and $h_\Delta := \frac{1}{2m}(\Delta p)^2 + \frac{m\omega^2}{2}(\Delta q)^2$. Since we know that q and p follow the classical trajectories, it must be the case that X_{h_Δ} is purely vertical. Let us see why this is the case. First, since $\frac{1}{i\hbar}[\hat{Q}^2, \hat{P}] = 2\hat{Q}$, if we take expectation values and apply Lemma 3.6, we obtain

$$X_p\big((\Delta q)^2\big) = \big\{(\Delta q)^2, p\big\} = 0. \tag{5.14}$$

Similarly, we find that if f_r and f_s are any two of q or p (or, in the case in which there are more basic variables, any two of the q_i and p_i),

$$X_{f_r}\big(\Delta f_s\big) = 0. \tag{5.15}$$

Therefore, the uncertainties in the basic variables are constant on any horizontal section! As a consequence, X_{h_Δ} is purely vertical.

If $s : \Gamma \to \mathcal{P}$ is any horizontal section, the pull-back s^*h is exactly the classical Hamiltonian function, up to the constant uncertainty term. This is essentially the reason that the evolution of the basic quantum observables agrees with the classical evolution. However, the quantum evolution is far from classical in that it does not generally preserve the horizontal sections. We are lead to ask if there is a horizontal section which is preserved by the Hamiltonian evolution; i.e., is there a section $s_0 : \Gamma \to \mathcal{P}$ on which X_{h_Δ} vanishes everywhere? If so, this would provide a preferred embedding of the classical phase space into $\mathcal{P}$.

Notice that X_h will be purely horizontal at $x \in \mathcal{P}$ if the uncertainty term h_Δ attains a local extremum there. It seems a rather natural guess that this will be the case at some state which saturates the uncertainty relation between q and p. To see that this is indeed the case, it is convenient to write h_Δ as

$$\frac{1}{\hbar\omega}h_\Delta = \left[\sqrt{\frac{m\omega}{2\hbar}}\Delta q - \sqrt{\frac{1}{2m\hbar\omega}}\Delta p\right]^2 + \frac{1}{\hbar}\Delta p\Delta q \;\;\geq\;\; \frac{1}{2}, \tag{5.16}$$

where the inequality is, of course, due to Heisenberg. Therefore, if we are to find a state x at which $h_\Delta(x) = \frac{\hbar\omega}{2}$ we will have found a state at which h_Δ attains a global minimum, and hence a state at which the flow is purely horizontal. This is easy, since a state at which $(\Delta q)^2 = \frac{\hbar}{2m\omega}$ and $(\Delta p)^2 = \frac{m\hbar\omega}{2}$ is easily seen to do the trick! Since the only states which saturate Heisenberg's inequality are the Gaussians, and those with the particular values of the uncertainties that we have just chosen are the only ones at which h_Δ is minimized, there is only a single horizontal section which is preserved by the Hamiltonian evolution. This is the preferred embedding of the classical phase space we have been seeking; it is the section, on which the quantum evolution is the "most classical". It is easy to see that this result holds for the higher dimensional oscillator as well.

5.3.2 Generalized coherent states

The reader who is familiar with the notion of coherent states may be wondering how they fit into the geometric description of quantum mechanics. First of all, as is pointed out in [48, 8], the generalized coherent state spaces (this notion is explained below) is endowed with a natural geometric structure. How is the geometry of the coherent state spaces related to the geometry of the quantum kinematics? Secondly, the standard space of coherent states of the harmonic oscillator may be specified by a submanifold of the Hilbert space, each point of which saturates the Heisenberg uncertainty relation in exactly the same way as does the special horizontal section discovered above. The (projection to $\mathcal{P}$ of the) standard coherent state space is exactly the preferred horizontal section! Let us consider this in more detail.

While we will consider only the oscillator, the notion of the coherent state is not particular to this case. There are a number of different constructions of the coherent state space. We will use the attractive, and fairly general, approach introduced by Perelomov [2] and Gilmore [3] (and indicated much earlier by Klauder [49]). Perelomov's *generalized coherent states* are obtained by the action of the dynamical group of the system on some fiducial element of the Hilbert space (see, e.g., Ref. [8]). The dynamical group of the oscillator—the Heisenberg-Weyl group, $\mathcal{W}_1$—is obtained by exponentiating the Lie algebra generated by $\hat{Q}$, $\hat{P}$, and $\mathbf{1}$. Let $t \in \mathbb{R}$ and let q' and p' be arbitrary real numbers with dimensions of length and momentum, respectively. Then

$$W(t,q',p') := \exp\left[-\mathrm{i}t\mathbf{1} - \frac{\mathrm{i}}{\hbar}\left(q'\hat{P} - p'\hat{Q}\right)\right] = \mathrm{e}^{-\mathrm{i}t}\,D'(q',p')$$

is the generic element of $\mathcal{W}_1$, where $D'(q',p') := \exp\left[-\frac{\mathrm{i}}{\hbar}\left(q'\hat{P} - p'\hat{Q}\right)\right]$. Recall that elements of $\mathcal{W}_1$ satisfy the 'commutation relations'

$$\begin{aligned} W(t_1,q_1',p_1')\,W(t_2,q_2',p_2') &= \exp\left[\frac{1}{\mathrm{i}\hbar}\left(q_1'p_2' - q_2'p_1'\right)\right] W(t_2,q_2',p_2')W(t_1,q_1',p_1') \\ &= W\left(t_1+t_2+\frac{1}{2\hbar}\left(q_1'p_2' - q_2'p_1'\right),\ q_1+q_2,\ p_1+p_2\right). \end{aligned} \tag{5.17}$$

Fix an arbitrary element $\Psi_0 \in \mathcal{H}$ of unit norm. Given any element W of the Heisenberg-Weyl group, the vector $W\Psi_0$ determines a coherent state (by definition). However, since $W(t,q',p')\Psi_0$

and $W(s,q',p')\Psi_0$ differ by a phase they determine the same physical state. One often chooses to consider only those elements of the form

$$\Psi_{(q',p')} := W(0,q',p')\Psi_0 = D'\big(q',p'\big)\Psi_0; \tag{5.18}$$

these elements of the Hilbert space are in one-to-one correspondence with the generalized coherent states.

Calculations are a bit easier if one introduces the complex coordinate $z := \big(\frac{m\omega}{2\hbar}\big)^{\frac{1}{2}} q' + \mathrm{i}\big(\frac{1}{2m\hbar\omega}\big)^{\frac{1}{2}} p'$ and defines the creation and annihilation operators,

$$\hat{A} := \sqrt{\frac{m\omega}{2\hbar}}\hat{Q} + \mathrm{i}\sqrt{\frac{1}{2m\hbar\omega}}\hat{P}, \qquad \hat{A}^\dagger := \sqrt{\frac{m\omega}{2\hbar}}\hat{Q} - \mathrm{i}\sqrt{\frac{1}{2m\hbar\omega}}\hat{P}.$$

Then $D(z) = D'\big(q',p'\big) = \exp\big[z\hat{A}^\dagger - \overline{z}\hat{A}\big]$. The coherent states are typically labeled by the complex parameter z:

$$\Psi_z = \Psi_{(q',p')} = D(z)\Psi_0. \tag{5.19}$$

It is easy to see that $D^\dagger(z)\hat{A}D(z) = \hat{A} + z$. Using this fact, one makes a few interesting observations. First, let us find the expectation values of $\hat{Q}$ and $\hat{P}$ at the coherent state Ψ_z:

$$\begin{aligned}\big\langle \Psi_z,\, \hat{Q}\Psi_z\big\rangle &= \big\langle \Psi_0,\, D^\dagger(z)\hat{Q}D(z)\Psi_0\big\rangle \\ &= \sqrt{\frac{\hbar}{2m\omega}}\,\big\langle \Psi_0,\, D^\dagger(z)\big(\hat{A}+\hat{A}^\dagger\big)D(z)\Psi_0\big\rangle \\ &= \big\langle \Psi_0,\, \hat{Q}\Psi_0\big\rangle + q'.\end{aligned} \tag{5.20}$$

Similarly, $\big\langle \Psi_z,\, \hat{P}\Psi_z\big\rangle = \big\langle \Psi_0,\, \hat{P}\Psi_0\big\rangle + p'$. The "displacement operator" $D(q',p')$ acts as a translation on the space labeled by the parameters q' and p'. The next observation that one typically makes is that the uncertainties in $\hat{Q}$ and $\hat{P}$ are *constant* on the entire coherent state space; similar to what was done in Eqs. (5.20),

$$\big(\Delta\hat{Q}\big)\big|_{\Psi_z} = \big(\Delta\hat{Q}\big)\big|_{\Psi_0} \quad \text{and} \quad \big(\Delta\hat{P}\big)\big|_{\Psi_z} = \big(\Delta\hat{P}\big)\big|_{\Psi_0}. \tag{5.21}$$

Having chosen the fiducial element Ψ_0, and a convention by which to fix the phase of the coherent states, one then obtains a submanifold of the Hilbert space on which the uncertainties of the basic operators are constant. Note that the construction goes through in the higher-dimensional case as well (one just adds more $\hat{Q}$s and $\hat{P}$s); the Maxwell field being the typical application. Note also that if one chooses, for Ψ_0, the ground state of the oscillator, the standard coherent states are obtained, but this is a special case of Perelomov's construction. The advantage of his generalization lies in its applicability to more general systems.

Of course, the coherent state space just described defines a submanifold of the projective Hilbert space. Notice that from the definition (5.18),

$$\begin{aligned}\frac{\partial}{\partial q'}\Psi_{(q',p')} &= \frac{1}{\mathrm{i}\hbar}\hat{P}\Psi_{(q',p')} = X_P\big|_{\Psi_{(q',p')}}, \\ \frac{\partial}{\partial p'}\Psi_{(q',p')} &= -\frac{1}{\mathrm{i}\hbar}\hat{Q}\Psi_{(q',p')} = -X_Q\big|_{\Psi_{(q',p')}}.\end{aligned} \tag{5.22}$$

Projecting these expressions to $\mathcal{P}$, we see that the submanifold of coherent states is precisely the horizontal section through $\pi(\Psi_0)$; in the previous section, we accidentally constructed the coherent state spaces by entertaining ideas which were naturally presented by the geometry of the quantum phase space!

5.3.3 Semi-classical dynamics

We have, to some extent, considered the relationship between classical and quantum dynamics from the point of view of Ehrenfest's theorem. The second approach which is often implemented involves a heuristic relationship between quantum mechanics and Hamilton-Jacobi theory, and is related to the WKB approximation. Our goal in this section is to understand the approximate WKB dynamics from a geometric point of view. We discuss only two results. First, we find that the WKB equation actually corresponds to a Hamiltonian vector field on the projective space; it is then described by a dynamics of the Weinberg type. Second, we find a fairly interesting statement of the validity condition of the WKB approximation which, to our knowledge, has not been discussed in the literature.

Let us consider the quantum theory of a non-relativistic particle moving in $\mathbb{R}^n$ under the influence of a conservative force. The wave function is then assumed to satisfy the Schrödinger equation

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi(x,t) = \left(-\frac{\hbar^2}{2m}\Delta + V(x)\right)\Psi(x,t), \tag{5.23}$$

where Δ is the Laplace operator on $\mathbb{R}^n$. Keeping in spirit with our previous approach, let us decompose the state vector into real and imaginary parts; we define fields ϕ and π via

$$\Psi =: \frac{1}{\sqrt{2\hbar}}\left(\phi + \mathrm{i}\pi\right).$$

One should note that, in terms of these fields, the symplectic structure and metric introduced in §2.3 assume the form

$$G\big((\phi_1,\pi_1),\ (\phi_2,\pi_2)\big) = \int \mathrm{d}^n x\,\big[\phi_1(x)\phi_2(x) + \pi_1(x)\pi_2(x)\big],$$
$$\Omega\big((\phi_1,\pi_1),\ (\phi_2,\pi_2)\big) = \int \mathrm{d}^n x\,\big[\phi_1(x)\pi_2(x) - \phi_2(x)\pi_1(x)\big].$$

The second equation looks quite familiar from classical field theory. The ϕ and π may be viewed as the 'field' and 'momentum density'; they are canonically related. Indeed, one may easily calculate the expectation function of the above Hamiltonian operator,

$$H(\phi,\pi) = \frac{1}{2\hbar}\int \mathrm{d}^n x\left[\frac{\hbar^2}{2m}\left((\vec{\partial}\phi)^2 + (\vec{\partial}\pi)^2\right) + V(x)\left(\phi^2(x) + \pi^2(x)\right)\right], \tag{5.24}$$

and verify that the fields evolve according to the canonical equations of motion:

$$\frac{\partial\phi}{\partial t} = \frac{\delta H}{\delta\pi} \qquad \text{and} \qquad \frac{\partial\pi}{\partial t} = -\frac{\delta H}{\delta\phi}. \tag{5.25}$$

Now, we are really interested in the relationship between this dynamics and the Hamilton-Jacobi equation. Recall that this relationship is typically introduced by writing $\Psi(x,t) = \sqrt{\rho(x,t)}\exp[\mathrm{i}S(x,t)/\hbar]$, where $\rho = \overline{\Psi}\Psi$. Inserting this into the Schrödinger equation immediately yields the equations

$$\frac{\partial S}{\partial t} + \frac{1}{2m}(\vec{\partial}S)^2 + V(x) = \frac{\hbar^2}{2m}\frac{\Delta\sqrt{\rho}}{\sqrt{\rho}}, \tag{5.26}$$

$$m\frac{\partial \rho}{\partial t} + \vec{\partial}\cdot(\rho\vec{\partial}S) = 0. \tag{5.27}$$

Notice that the second equation is a conservation equation, $\frac{\partial\rho}{\partial t} + \vec{\partial}\cdot\vec{J} = 0$, where $\vec{J} = \rho\vec{\partial}S/m$ is the probability current density. The observation that is so often made is the fact that Eq. (5.26) would look like the Hamilton-Jacobi equation if only the right hand side were absent. One should see Goldstein [50] or [17] for a discussion of the relevance of this observation. Our intention here is to examine the meaning of the dynamics that is obtained by simply dropping this "quantum mechanical correction" to the Hamilton-Jacobi equation; it seems fair to call this the *WKB evolution*.[1]

As a first step, one can express the Hamiltonian function of Eq. (5.24) in terms of the fields ρ and S:

$$H(\rho,S) = \int \mathrm{d}^n x \left[\frac{\hbar^2}{8m\rho(x)}\left(\vec{\partial}\rho\right)^2 + \frac{1}{2m}\rho(x)\left(\vec{\partial}S\right)^2 + \rho(x)V(x)\right]. \tag{5.28}$$

It may be seen that ρ and S are actually canonically related; it is easy to explicitly check that the above equations of motion are given by $\frac{\partial\rho}{\partial t} = \frac{\delta H}{\delta S}$ and $\frac{\partial S}{\partial t} = -\frac{\delta H}{\delta\rho}$. Next, let $H_\hbar$ equal the first term in the above expression for the Hamiltonian function, and put $H_{\mathrm{WKB}} := H - H_\hbar$. Then the Hamiltonian vector field corresponding to H_{WKB} corresponds exactly to the WKB evolution. Let us denote by X_{WKB} the Hamiltonian vector field (on $\mathcal{H}$) determined by H_{WKB}.

That the WKB evolution is generated by a Hamiltonian vector field is not obvious; it is even less obvious that it induces a Hamiltonian vector field on the projective space, but it is true. To see this, we need only find that X_{WKB} preserves the unit sphere S; this is guaranteed iff the Poisson bracket of H_{WKB} and the constraint function, $C = \int \mathrm{d}^n x\,\rho(x) - 1$, vanishes on S. Now, since $H_\hbar$ is independent of S, $\{H_\hbar, C\} = 0$. We know that $\{H, C\} = 0$ because H generates the Schrödinger flow, which preserves norms. Therefore, $\{H_{\mathrm{WKB}}, C\} = \{H - H_\hbar, C\} = 0$. One may wish to notice that this Poisson bracket vanishes *strongly*; the WKB evolution therefore preserves the norm of all state vectors (not just the unit vectors). The implication of the fact that the WKB evolution is given by a Hamiltonian vector field which preserves S is that it corresponds to dynamics of the Weinberg type; in particular, it defines a Hamiltonian vector field on the projective space.

This observation is the main point of this section. However, let us briefly ask for the condition which guarantees that the WKB evolution approximates the Schrödinger evolution; i.e., we ask for the condition of validity of the WKB *approximation*. Let us first calculate the Hamiltonian vector field determined by $H_\hbar$; we need the functional derivatives of $H_\hbar$ with respect to the fields ϕ and π.

[1] Actually, the term 'evolution' may not be fair after all. We have a differential equation, but there is no guarantee that there is a corresponding *flow*. This is, in general, a difficult question which we will ignore.

This is fairly straight-forward; one obtains

$$\frac{\delta H_\hbar}{\delta\phi} = -\frac{1}{8m\hbar\rho^2}\left[2\hbar\rho\big(\phi\Delta\phi + \pi\Delta\pi\big) + \big(\phi\vec{\partial}\pi - \pi\vec{\partial}\phi\big)^2\right]\phi, \tag{5.29}$$

and $\delta H_\hbar/\delta\pi$ is obtained by making the replacement $\phi \leftrightarrow \pi$ in the above equation. To understand the 'physical' meaning of Eq. (5.29) we have to make a preliminary observation. Let us write the expectation values of $\vec{P}$ and $\hat{P}^2$:

$$\begin{aligned} \big\langle \vec{P} \big\rangle &= \int \mathrm{d}^n x\; m\vec{J}(x), \qquad \vec{J} = \frac{1}{2}\big(\phi\vec{\partial}\pi - \pi\vec{\partial}\phi\big), \\ \big\langle \hat{P}^2 \big\rangle &= \int \mathrm{d}^n x\; K(x), \qquad K(x) := -\frac{\hbar}{2}\big[\phi\Delta\phi + \pi\Delta\pi\big]. \end{aligned} \tag{5.30}$$

Again, $\vec{J}$ is the probability current density. The quantities $m\vec{J}$ (the "density of momentum") and K (the "density of momentum-squared") are the terms that appear in the above expression of the relevant functional derivative.

In terms of these quantities, the functional derivatives expressed in Eq. (5.29) may be written as

$$\begin{aligned} \frac{\delta H_\hbar}{\delta\phi} &= -\frac{1}{2m\hbar\rho^2}\left[\big(m\vec{J}\big)^2 - \rho K\right]\phi, \\ \frac{\delta H_\hbar}{\delta\pi} &= -\frac{1}{2m\hbar\rho^2}\left[\big(m\vec{J}\big)^2 - \rho K\right]\pi, \end{aligned} \tag{5.31}$$

Therefore,

$$\mathrm{i}\hbar\left[\frac{\delta H_\hbar}{\delta\pi} - \mathrm{i}\frac{\delta H_\hbar}{\delta\phi}\right] = -\frac{1}{2m\rho^2}\left[\big(m\vec{J}\big)^2 - \rho K\right](\phi + \mathrm{i}\pi). \tag{5.32}$$

The WKB evolution is obtained by subtracting ($\frac{1}{\sqrt{2\hbar}}$ times) this term from the Schrödinger equation (c.f. Eq. (5.25)) ;

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi = \hat{H}\Psi + \frac{1}{2m\rho^2}\left[\big(m\vec{J}\big)^2 - \rho K\right]\Psi, \qquad \text{(WKB equation)}$$

where $\hat{H}$ is the Hamiltonian operator appearing in Eq. (5.23). The condition of validity for the WKB approximation is just that this additional term be small; i.e., the "density of momentum-squared" (weighted by the probability density) must nearly equal the square of the "density of momentum". While this condition has an intuitive flavor, it is a bit puzzling. It would be fun to examine this further, but we'll leave that for another day.

6

Summary and Conclusions

We have seen that quantum mechanics admits a geometric formulation which is intrinsic to the projective Hilbert space. The first key observation, made in Ch. 2, was that the Hermitian inner-product on $\mathcal{H}$ naturally defines a symplectic structure and a positive-definite inner-product (both of which are strongly non-degenerate). A Hilbert space is a real Kähler space. Secondly, the expectation value of a bounded, self-adjoint operator, as a function on the Hilbert space, contains all information about the operator itself. Recalling these facts have lead us to a complete formulation of quantum mechanics in terms of the true space of pure states—the projective Hilbert space $\mathcal{P}$. In this geometric description, the observables of the formalism are not represented by bounded, self-adjoint operators on $\mathcal{H}$, but by the associated real-valued functions on $\mathcal{P}$.

We found that many features of ordinary quantum mechanics are described by the mathematics of classical physics. A bounded, self-adjoint operator $\hat{F}$ defines a vector field on $\mathcal{P}$; integration of this vector field simply produces the (projection to $\mathcal{P}$ of the) unitary flow defined by $\hat{F}$. There is another vector field on $\mathcal{P}$ which is naturally defined by $\hat{F}$—the Hamiltonian vector field associated to the corresponding expectation value function. These two vector fields are identical. In particular, the Schrödinger equation assumes the form of Hamilton's canonical equation of motion! Moreover, the Poisson bracket of two expectation value functions is exactly the expectation value of ($-\mathrm{i}/\hbar$ times the) commutator of the respective operators. The Lie algebraic structure on the space of quantum observables, usually described by the commutator of linear operators, may then be described by a Poisson bracket, just as in classical mechanics. In this way, quantum mechanics may be viewed as an infinite-dimensional Hamiltonian system. As a result, one should expect quantum mechanical perturbation theory to be described just as one might do so for classical mechanics. In the same vein, the adiabatic theorems of classical and quantum mechanics presumably translate to exactly the same statements.

All of the mathematical features of classical mechanics are precisely implemented when the quantum physics is formulated in the geometric language. The quantum mechanical formalism is, however, much more structured than is classical mechanics. Complementing the 'classical' structure of the formulation are features which are particular to quantum mechanics. Although a Riemannian metric is often present in classical mechanics, it plays no *essential* role in the formulation of classical physics. On the other hand, Riemannian geometry is an *indispensable* feature of quantum mechanics.

In fact, as was seen in Ch. 3, the set of quantum observables consists of all real functions on $\mathcal{P}$ whose Hamiltonian vector fields preserve the metric; the observable functions are those which preserve *all* of the kinematic structure. In addition, the notion of uncertainty and the probabilistic aspects of quantum mechanics rely on the metric in a crucial way. For example, the uncertainty of an observable, as a function on the quantum phase space, is (up to an overall constant factor) the length of its Hamiltonian vector field. Intuitively, the faster the corresponding flow on the phase space, the less accurately one may measure the observable. The measurement process is also conveniently described by the metric; the relevant probabilities and spectral projections are directly related to geodesic distance. A particular result is the 'laziness' of the reduction process—as a result of measurement, a quantum state is more likely to collapse to a nearby state than a distant one.

Although the Riemannian structure is very important for the description of the physics, it is tempting to momentarily ignore the metric. The quantum phase space (along with the symplectic flow which defines the evolution) then defines an artificial classical theory. One is compelled to ask what would result from quantization of this fake classical theory. This question was answered in Ch. 4. We found that application of the geometric quantization procedure simply produces the original Hilbert space. Of all functions on $\mathcal{P}$, those which unambiguously give rise to 'quantum operators' on the Hilbert space are exactly the observable functions of the original quantum theory. Moreover, the operator obtained by quantizing any one of these functions is the operator which defined it in the first place! In this sense, quantization is a stable procedure. Upon closer examination, however, we noticed that there are actually (countably) many different descriptions of the quantum phase space, with our particular choice being, in a loose sense, the most natural one. So we quantized them all. The Hilbert spaces that result are simply the many particle subspaces of the symmetric Fock space. Again, the 'quantizable' functions are the observable ones. Further, let $f : \mathcal{P} \to \mathbb{R}$ be the observable function defined by the operator $\hat{F}$. We may construct two operators on the k-particle subspace—the natural action of $\hat{F}$, and that operator obtained by the k^{th} quantization of the observable function f. These two operators are identical. By quantizing the projective Hilbert space in *all* possible ways, we obtain the familiar Fock space and the associated second quantized observables. Second quantization really is second quantization.

One of the most attractive, and potentially useful, features of the geometric language is the separation of the mathematical structures relevant for the description of the unitary evolution and measurement process, respectively. While the symplectic structure describes those features which have analogues in classical mechanics, all of the purely quantum mechanical aspects of the formalism depend on the metric. The distinction between 'U' and 'R' is made explicit in terms of the mathematical structures employed in the geometric formulation. This fact suggests, for example, interesting approaches to the measurement problem which will be examined in the future.

As discussed in the Introduction, one may expect a resolution of the measurement problem to involve another important problem in theoretical physics—quantum gravity. It is partly for this reason that a framework in which to study non-linear modifications of quantum mechanics is desired. In Ch. 5 we made some preliminary investigations in this direction. In any attempt to study

generalizations of a physical framework, it is imperative that one characterize the original theory in terms of those structures which are being extended. In §5.1 we considered, as a radically generalized quantum kinematics, an arbitrary Kähler manifold $\mathcal{M}$. As suggested by ordinary quantum mechanics, we supposed the set of generalized observables to consist of all (smooth) functions on $\mathcal{M}$ whose Hamiltonian vector fields are Killing. The main result was the fact that if this set of observables is *maximal* and closed under the natural commutative operation (which stems from the Jordan product of quantum operators), then $\mathcal{M}$ is a manifold of constant holomorphic sectional curvature—a key property of the projective Hilbert space. While this result may be used to classify the projective Hilbert spaces of finite dimension, we can not go quite this far in the infinite-dimensional context. This indicates both a shortcoming and a possibly interesting result. Of course, we would like a complete classification of ordinary quantum mechanics. However, as suggested on page 72, if this classification is found to be quite subtle, it is likely that there exists a *minor* modification of ordinary quantum mechanics with physically interesting properties. Both aspects of this issue clearly deserve more attention than they have recieved here.

The sort of generalization outlined above was considered mainly for the purpose of understanding the structure which specifies quantum mechanics from all possible generalizations. Of course, there are many other sorts of generalized mechanics' that one might invent. For example, one may work with the standard kinematic structures but generalize only the dynamics; a time evolution which preserves the symplectic structure, but not necessarily the metric, seems an obvious choice to consider. The results of §5.2 suggest that many (if not all) of the non-linear modifications to the Schrödinger equation that have been considered by previous researchers are of just this form. One can go a step further and extend the definition of the observables as well. In particular, it may seem rather natural (though not physically justified) to consider the space of *all* functions on $\mathcal{P}$ as 'observables'; this is the meaning of Weinberg's proposal.

We have touched upon the idea of generalized quantum mechanics only very briefly, and for good reason; it is a difficult problem. Quantum mechanics is extremely rigid in that consistency of the probabilistic interpretation seems deeply entwined with the particular structure of the projective Hilbert space. In this regard, we have not gone quite far enough with the geometric description of quantum mechanics. In particular, the fact that the probabilities, as defined in the geometric language, add to one is not at all obvious from the formalism. This is the major barrier to obtaining an understanding of the conditions which guarantee that a generalized mechanics admits a consistent probabilistic formulation. For example, it would be desirable to have a clearer description of the eigenmanifolds (or the image of a spectral projection, in the case of the continuous spectrum) described in Ch. 3. Of course, to fully develop a generalized mechanics, one must extend the standard ideas employed in spectral analysis. It might be useful to note that Morse theory[1] could be a very useful tool with which to pursue "generalized spectral analysis".

Also related to the issue of measurement is an important feature of quantum mechanics that our presentation is noticeably lacking. The outcome of any measurement is due to the interaction of a

[1] This was suggested by Joseph Samuel during personal conversation.

system with some measurement apparatus. We are accustomed to modeling this situation by the formation of the tensor product of the Hilbert spaces of the two constituents. Unfortunately, we do not know of a simple geometric description of the projective Hilbert space corresponding to such a tensor product. Therefore, issues involving spin-statistics have not been seen to admit a natural geometric formulation. While we have not considered this question in detail, it is not unlikely that this issue can be resolved in a fairly simple manner. In the finite-dimensional regime, a sub-Hilbert space corresponds to a certain type of submanifold of the projective space known as an algebraic variety; therefore, there is some well-developed mathematics with which to study this question.[2]

Generalizations of quantum mechanics provided one motivation for examining the geometric properties of quantum mechanics. A further motivation, which was also considered in Ch. 5, is the relationship between classical and quantum mechanics. For the special case in which the classical phase space is linear, we found an interesting description of the generalized coherent state spaces. First, we discovered that one may view the quantum phase space as a fibre bundle over the classical one. The quantum symplectic structure then allowed a natural definition of horizontal spaces which are, in fact, integrable. Two interesting results follow. First is a construction of the classical symplectic structure from the quantum symplectic structure; this is a geometric statement of the kinematic classical limit. Secondly, the horizontal sections define numerous special embeddings of the classical phase space in the quantum phase space. As we found, these embeddings are nicely described in terms of the coherent states of the harmonic oscillator. Finally, we considered semi-classical dynamics from the point of view of the WKB approximation. The WKB evolution is typically described in a way which relies heavily on fact that $\mathcal{H}$ is a linear space. The fact that the WKB evolution defines a Hamiltonian vector field on the projective space is far from obvious, and quite interesting. It suggests that even dynamical aspects of the classical limit may be explored in purely geometric terms. These latter observations are preliminary and it would be interesting to follow them to their logical conclusions.

The subject of non-linear quantum mechanics suggests a very interesting question. The description of chaotic classical systems relies on the non-linearity of the Hamiltonian equation of motion. Since the evolution of an ordinary quantum mechanical system is, by definition, linear, no ordinary quantum system is chaotic in the classical sense. The notion of quantum chaos is therefore not defined in terms of the unitary (i.e., Hamiltonian) evolution. Rather, one must resort to considering, for example, the dependence of the energy eigenvalues on non-dynamical coupling coefficients that appear in the Hamiltonian. One might suggest that this seemingly odd practice describes the chaotic behavior of the 'R-process' rather than the 'U-process'. The author, however, knows of no precise implementation of this idea. Alternatively, however, a non-linear modification of quantum mechanics (as might be employed to model the measurement process) *does* suggest the possibility of a link between classical and quantum chaos.

The geometric desciption is not only attractive, but suggests many interesting questions which we have only begun to consider.

[2]Thanks to Lane Hughston for pointing this out.

A

Fibre Bundles

The fibre bundle is the natural setting in which to discuss gauge theory, covariant differentiation of tensor fields on a manifold, etc. The purpose of this appendix is primarily to establish our notational conventions. The summary presented here is quite brief; only those notions essential for the development of the ideas in the main text are presented. For a detailed account see, for example, Chs. 9 and 10 of [30].

The simplest example of a fibre bundle is the direct product $B \times F$, of two manifolds. The general case is a slight abstraction; a manifold[1] E, which is *locally* a direct product. This is somewhat analogous to the definition of a differentiable manifold, which locally "looks like" a vector space, but does not have a global linear structure.

More precisely, a fibre bundle is a manifold E, along with the following structures:

i. A manifold B, called the *base manifold*.

ii. A manifold F, called the *fibre manifold*.

iii. A *projection mapping* $\pi : E \to B$ which is surjective.

iv. A Lie group G, called the *structure group* or *gauge group*, which acts on F on the left.

v. An open cover $\{U_i\}$ of B with associated diffeomorphisms $\phi_i : U_i \times F \to \pi^{-1}(U_i)$ such that $\pi \circ \phi_i(p, f) = p$. The ϕ_i are called *local trivializations*. The set $\pi^{-1}(p)$ is then diffeomorphic to F, and is called the *fibre at* p (or *above* p).

vi. On the overlap $U_i \cap U_j \neq \emptyset$, we define the *transition function* $g_{ij} : U_i \cap U_j \to G$ by the relation $\phi_j(p, f) = \phi_i(p, g_{ij} f)$.

Note that we require that the transition functions be smooth. Of course, the g_{ij} must satisfy some consistency requirements. For example, on $U_i \cap U_j \cap U_k$, we must have $g_{ij} g_{jk} = g_{ik}$. Further, it is convenient to assume that our set of local trivializations is maximal in the sense that if U is an open

[1] The manifolds below are all assumed to be equipped with differential structures; i.e., we consider only *differentiable* fibre bundles.

subset of B and $\phi : U \times F \to \pi^{-1}(U)$ is compatible with $\{(U_i, \phi_i)\}$ in the sense suggested above, then $(U, \phi) \in \{(U_i, \phi_i)\}$.

As a notational convenience, we follow tradition in denoting the above bundle as $E \xrightarrow{\pi} B$. The particular Lie group G and fibre space F are usually clear from the context.

A *section* of the fibre bundle $E \xrightarrow{\pi} B$ is a smooth map $s : B \to E$ such that $\pi \circ s(p) = p$. To each point $p \in B$, the section associates an element $s(p)$ in the fibre above p. We also consider *local* sections; i.e., mappings $s : U \subset B \to E$ defined only on open subsets of B.

A common example is the tangent bundle $E = T^*B$ of a manifold B. An element of T^*B is a vector in the tangent space at $p \in B$. Sections of T^*B are vector fields on B.

Principal Bundles

A fibre bundle $S \xrightarrow{\pi} P$ for which the fibre space coincides with the gauge group G, is called a *principal bundle*. Since G acts on itself in the obvious way, we may define a right action of the group on the bundle. Given a local trivialization $\phi_i : U_i \times G \to \pi^{-1}(U_i) \subset S$ and an element $g \in G$, the right action of g is defined locally by $\phi_i(p, u) \cdot g := \phi_i(p, ug)$. Since the left action of G on itself commutes with the right action, this provides us with a well-defined right action of G on S. Note that the left action of G on the fibre space does *not* allow us to define a left action of the structure group on the generic fibre bundle.

This allows us to set up a correspondence between local sections and local trivializations. Given (U_i, ϕ_i), we may define the local section $\sigma_i : U_i \to \pi^{-1}(U_i)$ via

$$\phi_i(p, u) =: \sigma_i(p) \cdot u. \tag{A.1}$$

Conversely, to a local section σ_i is associated the local trivialization determined by the above equation. If (U_i, ϕ_i) and (U_j, ϕ_j) are related by the transition function g_{ij} as in item (vi) above, then the canonical sections are related by the right action

$$\sigma_j(p) = \sigma_i(p) \cdot g_{ij}(p) \quad \forall\, p \in U_i \cap U_j. \tag{A.2}$$

Connections on Principal Fibre Bundles

On a generic fibre bundle, there is a natural notion of *verticality* of tangent vectors on E. A vector $v \in T_xE$ is vertical if $\pi_* v = 0$; i.e., if v is in the direction of the fibre containing x. As a particularly useful example, consider a principal bundle $S \xrightarrow{\pi} P$. Let $x \in S$, and suppose L is an element of the Lie algebra of the structure group G. We may then define a vertical vector at x by use of the right action of G on S:

$$L^\sharp|_x := \frac{d}{dt}[x \cdot \exp(Lt)]_{t=0}.$$

In this way, one obtains from L a vector field on S, called the *fundamental vector field generated by* L. This operation is available any time one has a group action on a manifold. Notice that we may make use of the above definition for any Lie algebra valued function L on the bundle manifold.

A principal bundle does not, however, come with a prescribed notion of *horizontality*. Denote by V_x the vertical subspace of T_xS as defined above. A *connection* is the specification of a horizontal subspace $H_x \subset T_xS$ at each point $x \in S$ such that:

i. $T_xS = V_x \oplus H_x \quad \forall \quad x \in S$.

ii. $R_{g*}(H_x) = H_{R_g(x)}$, where $R_g(x) := x \cdot g$; i.e., H_x is preserved by the right action of the gauge group.

In practice, a connection is typically specified by the use of a Lie algebra valued one-form $\mathcal{A}$ on S. The vertical part of a vector $v \in T_xS$ is then given by the projection $v \mapsto (\mathcal{A}(v))^\sharp$. The kernel of $\mathcal{A}|_x$ then defines the horizontal space at x. For consistency with the above definition, $\mathcal{A}$ must satisfy the requirements:

i. $\mathcal{A}(L^\sharp) = L$ for each Lie algebra element, L.

ii. $R_g^*\mathcal{A} = \mathrm{Ad}_{g^{-1}}\mathcal{A}$.

In item (ii), $\mathrm{Ad}_{g^{-1}}$ is the adjoint mapping on the Lie algebra defined by the group element g^{-1}; recall that for matrix groups, this takes the form $\mathrm{Ad}_{g^{-1}} : L \mapsto g^{-1}Lg$.

More familiar to physicists is the local description of the connection form. Given a local trivialization (U_i, ϕ_i), or equivalently the canonical section $\sigma_i : U_i \to \pi^{-1}(U_i)$, one may obtain a Lie algebra valued one-form on U_i by pulling back via σ_i;

$$A_i := {\sigma_i}^*\mathcal{A}.$$

Suppose (U_j, ϕ_j) is another local trivialization such that $U_i \cap U_j \neq \emptyset$, and let $\sigma_j : U_j \to \pi^{-1}(U_j)$ be the associated local section. With our conventions, the local expression for the connection form transforms as

$$A_j = {g_{ij}}^{-1}A_i g_{ij} + {g_{ij}}^{-1}dg_{ij} \tag{A.3}$$

on $U_i \cap U_j$.

A connection allows one to lift a curve $p(t)$ on $\mathcal{P}$ to a curve on S, in a special way. First, a lift $\Psi(t)$ of $p(t)$ is a curve for which $\pi \circ \Psi(t) = p(t)$. Given a connection on S, it is natural to consider *horizontal lifts*; i.e., those lifts whose tangents remain horizontal. Given a point $\Psi(0)$ in the fibre above $p(0)$, there exists a unique horizontal lift of $p(t)$. Denote this horizontal lift by $\tilde{p}(t)$. One can show that the horizontal lift through $\Psi(0) \cdot g$ is simply $\tilde{p}(t) \cdot g$, for all t; two horizontal lifts are related by *constant* right translation along the fibres. It is in this way that the notion of the holonomy arises. While we will not be concerned with the holonomy group, the notion of the horizontal lift will become useful shortly.

Associated Bundles

Suppose we have a linear representation[2] of G on a vector space V. To each element $g \in G$ is associated the linear map $\rho_g : V \to V$ To avoid the clutter of additional notation, we will simply write $g : v \mapsto gv$ rather than $\rho_g : v \mapsto \rho_g(v)$. Given a principal bundle $S \xrightarrow{\pi} \mathcal{P}$, we may define a vector bundle over $\mathcal{P}$ whose typical fibre is V as follows.

There is a natural action of G on the product $S \times V$, given by $g : (\Psi, v) \mapsto (\Psi g, g^{-1}v)$. The associated bundle manifold is defined by dividing $S \times V$ by this group action. In other words, one defines the equivalence relation

$$(\Psi, v) \sim (\Psi g, g^{-1}v),$$

and puts $E := (S \times V)/\sim$. Denote by $[\Psi, v]$ the equivalence class containing (Ψ, v). Note that $[\Psi, gv] = [\Psi g, v]$. The projection mapping is defined as

$$\pi_E : E \to \mathcal{P}, \qquad \pi_E : [\Psi, v] \mapsto \pi(\Psi). \tag{A.4}$$

That π_E is, indeed, a projection is obvious since $\pi(\Psi) = \pi(\Psi g)\ \forall g \in G$.

The bundle structure of $E \xrightarrow{\pi_E} \mathcal{P}$ is directly inherited from that of $S \xrightarrow{\pi} \mathcal{P}$ in the following manner. Given a local trivialization (U_i, ϕ_i) of the principal bundle, define the mapping

$$\tilde{\phi}_i : U_i \times V \to \pi_E^{-1}(U_i), \quad \tilde{\phi}_i : (p,\ v) \mapsto [\phi_i(p, e),\ v] = [\phi_i(p, g),\ g^{-1}v], \tag{A.5}$$

where e is the identity element of G. Recall that $\phi_i(p, e)$ is the local section $\sigma_i : U_i \to \pi^{-1}(U_i)$, canonically associated to the trivialization (U_i, ϕ_i). Finally, let $(U_i, \tilde{\phi}_i)$ and $(U_j, \tilde{\phi}_j)$ be local trivializations, as just defined. The transition functions for the associated bundle coincide with those for the principal bundle:

$$\begin{aligned} \tilde{\phi}_j(p,\ v) &= [\phi_j(p, e),\ v] = [\phi_i(p, g_{ij}(p)e),\ v] = [\phi_i(p, e) \cdot g_{ij}(p),\ v] \\ &= [\phi_i(p, e),\ g_{ij}(p)v] = \tilde{\phi}_i(p,\ g_{ij}(p)v). \end{aligned} \tag{A.6}$$

The associated bundle 'twists' in the same way as the principal bundle. Recall that a principal bundle is trivial if and only if it admits a global cross section. In general, there exists no global section of a principal bundle. However, there are always plenty of global sections of the associated vector bundle. Although there is no canonical identification of the fibres, each one has one preferred element—the zero element. The zero section of E, and deformations of it, are global.

Covariant differentiation

Let $E \xrightarrow{\pi_E} \mathcal{P}$ be a vector bundle associated to the principal bundle $S \xrightarrow{\pi} \mathcal{P}$, as just described. Above each point $p \in \mathcal{P}$ is a vector space; but there is no natural identification of one fibre with

[2] We don't need to restrict ourselves to linear representations here. All that follows is valid as long as we have a group action of G on a manifold. For simplicity, we assume this manifold is a vector space and that the group action is linear.

another. However, if one has a connection on the principal bundle, there is a natural way to parallel transport a vector along a curve, just as one does for the tangent bundle, for example. Naturally, this is defined in terms of parallel transport of points on the principal bundle, i.e., horizontal lifts. Let $p(t)$ be a curve on the base space $\mathcal{P}$, and let $[\Psi, v] \in E$ be an element of the fibre above $p(0)$. The horizontal lift of $p(t)$, through $[\Psi, v]$ is given by the curve $[\tilde{p}(t), v]$, where $\tilde{p}(t)$ is, of course, the horizontal lift to S, and v is held constant. Let $[\Psi(t), v(t)] = [\Psi(t)g(t),\ g^{-1}(t)v(t)]$ be any curve in E. We may always choose a 'gauge' for which $\Psi(t)g(t)$ defines a horizontal curve in $\mathcal{P}$. This can be accomplished, for example, by calculating a path-ordered exponential. The curve $[\Psi(t), v(t)]$ is horizontal if $g^{-1}(t)v(t)$ is constant, where $g(t)$ defines such a gauge.

Using this notion, it is now easy to introduce the idea of the covariant derivative of a section $s(p) = [\Psi(p), v(p)]$. Let X be an element of the tangent space $T_q\mathcal{P}$. X may be described by a curve $p(t)$, through q. We may parallel transport $s(0)$ along this curve, and make a comparison with $s \circ p(t)$ for infinitesimal values of t. This is the role of the covariant derivative; it measures the lack of horizontality of a section in any direction one chooses. One proceeds as follows. As in the previous paragraph, we can write $s(p(t)) = [\tilde{p}(t), \tilde{v}(t)]$, where $\tilde{p}(t)$ is horizontal. The covariant derivative of s in the direction of X is defined so that it vanishes in the event that $s \circ p$ is horizontal at q:

$$\mathcal{D}_X s := \left[\tilde{p}(0),\, \frac{\mathrm{d}}{\mathrm{d}t}\tilde{v}(p(t))\bigg|_{t=0}\right]. \tag{A.7}$$

By doing this at each point of $\mathcal{P}$, one may take the covariant derivative of a section with respect to a vector *field* on $\mathcal{P}$; the result is another section of E. If X is such a vector field, we write

$$\mathcal{D}_X s = X^\alpha \mathcal{D}_\alpha s. \tag{A.8}$$

The covariant derivative operator $\mathcal{D}_\alpha$ satisfies all of the properties one would expect of a derivative operator. In particular, it is linear and satisfies the Leibnitz rule with respect to multiplication of sections by functions.

Again, many physicists are more familiar with the local expression of the covariant derivative. Given any local trivialization $(U_i, \tilde{\phi}_i)$, a section $s : \mathcal{P} \to E$ may be written, locally, as $s(p) = [\sigma_i(p), z(p)]$, where $\sigma_i : U_i \to \pi^{-1}(U_i)$ is the local section canonically associated to ϕ_i. Locally, a section of E looks like a V-valued function. Let A_i be the local connection form on U_i, as in Eq. (A.3). The covariant derivative then assumes the local form

$$\mathcal{D}_X[\sigma_i, z] = [\sigma_i,\ X^\alpha \partial_\alpha z + X^\alpha A_{i\alpha}(z)]\,,$$

where the Lie algebra element $X^\alpha A_{i\alpha}$ acts on the vector z by the representation of G on V. This is usually written in terms of the vector component alone:

$$\mathcal{D}_\alpha z^I = \partial_\alpha z^I + A_\alpha{}^I{}_J z^J, \tag{A.9}$$

where the capital latin indices indicate the representation space, and the subscript 'i', labeling the trivialization, has been dropped for convenience.

B

Geometric Quantization

The method of geometric quantization is an attractive approach to the problem of quantization. As in the method of canonical quantization, it focuses on the geometric structure of the classical phase space; one tries to replace Poisson brackets of classical observables by ($1/\mathrm{i}\hbar$ times) the commutator of the corresponding quantum operators. However, it adds precision to the canonical procedure by specifying a particular way to "hunt down" the appropriate Hilbert space. When adopting the canonical approach, one often constructs the quantum Hilbert space from the space of square-integrable functions (or, better yet, half-densities) on the configuration space. If the classical phase space is not a cotangent bundle, there is no naturally defined notion of the configuration space; geometric quantization gives one a general approach, even in this situation. The idea is to introduce a structure which, in a sense, locally describes the phase space as a cotangent bundle. This appendix presents a lightning review of geometric quantization. For slightly more detailed review see [51]. For a complete treatment refer to [41] or [43].

Pre-quantization

The classical system, whose quantum mechanical description we would like to find, is given by some symplectic manifold (Γ, Ω). We will not touch on the (typically difficult) issue of quantization of the dynamics, so we will not include the Hamiltonian as part of the given structure. The pre-quantization process involves the construction of an algebra of self-adjoint operators which faithfully represents the entire Poisson algebra. Woodhouse [43, Ch. 5] beautifully motivates us to choose, for the pre-quantum Hilbert space, a complex line bundle over the phase space. This line bundle is associated to a principal $U(1)$-bundle, $S \xrightarrow{\pi} \Gamma$, equipped with a connection, $\mathcal{A}$. Given the principal bundle, the associated line bundle is obtained, as in Appendix A, with the vector space $V = \mathbb{C}$. It is absolutely essential that the curvature of the connection be related to the classical symplectic structure as[1]

$$\mathrm{d}\mathcal{A} = -\frac{1}{\hbar}\pi^*\Omega. \tag{B.1}$$

[1] We will use the language, typical in physics, which represents the Lie algebra of $U(1)$ with the real line.

The existence of the pre-quantum $U(1)$-bundle is not guaranteed. Such a bundle will exist if and only if $\frac{1}{\hbar}\Omega$ defines an integral de Rham cohomology class. This *pre-quantization condition* is equivalent to the statement that the integral of Ω over any closed, oriented 2-surface is an integral multiple of $2\pi\hbar$. Note that for the standard example, in which Γ is a cotangent bundle and Ω is the fundamental two-form, this condition is trivially satisfied, since Ω is exact.

Suppose the pre-quantization condition is satisfied, and that we have found the pre-quantum bundle $S \xrightarrow{\pi} \Gamma$, and its associated line bundle $E \xrightarrow{\pi_E} \Gamma$. Pre-quantum states are to be sections of E. Thus, a pre-quantum state *locally* looks like a complex-valued function on the phase space; this does not hold globally since there will not, in general, exist a global trivialization. Suppose $(U_i, \tilde{\phi}_i)$ and $(U_j, \tilde{\phi}_j)$ are two local trivializations of the line bundle. As described in Appendix A, these trivializations define two complex-valued functions $z_i : U_i \to \mathbb{C}$ and $z_j : U_j \to \mathbb{C}$. On the overlap $U_i \cap U_j$ these local functions are related as (c.f. Eq. (A.6))

$$z_j(p) = g_{ij}(p)\, z_i(p), \tag{B.2}$$

where g_{ij} is the transition function defined for the $U(1)$-bundle. Since these local functions just differ by a phase, given two sections ξ_1 and ξ_2, the integral

$$\langle \xi_1,\, \xi_2 \rangle_{pre} = \int_\Gamma \mathrm{d}v\; \overline{z_1} z_2 \tag{B.3}$$

is well-defined. Here, we are assuming that Γ is of (finite) dimension $2n$, so that we can define the volume form $\mathrm{d}v = \wedge^n \left(\frac{2\pi}{\hbar}\Omega\right)$. Denote by $\mathcal{H}_{pre}$ the (completion of the) set of all square-integrable sections of E. This is the pre-quantum Hilbert space.

To every classical observable $f : \Gamma \to \mathbb{R}$ is associated a formally self-adjoint (i.e., symmetric) operator on $\mathcal{H}_{pre}$ by the definition

$$\hat{O}_f \xi := -\mathrm{i}\hbar {X_f}^\alpha \mathcal{D}_\alpha \xi + f\xi,$$

where $\mathcal{D}$ is the covariant derivative determined by the connection $\mathcal{A}$. The second term is included so that the map $f \mapsto \hat{O}_f$ provides a faithful representation. In terms of a local trivialization $(U, \tilde{\phi})$ of the line bundle, a section may be written as $\xi(p) = [\sigma(p), z(p)]$, where $\sigma : U \to \pi^{-1}(U)$ is the canonical section. The pre-quantum operator above then assumes the local form (c.f. Eq. (A.9))

$$\hat{O}_f z = -\mathrm{i}\hbar {X_f}^\alpha (\partial_\alpha - iA_\alpha) z + fz = -\mathrm{i}\hbar X_f(z) + [f - \hbar A(X_f)] z. \tag{B.4}$$

One can show that the pre-quantum operators are symmetric. Further, the map $f \mapsto \hat{O}_f$ takes Poisson brackets to commutators; for *any* two classical observables f and g,

$$\hat{O}_{\{f,g\}} = \frac{1}{\mathrm{i}\hbar}\left[\hat{O}_f,\, \hat{O}_g\right]. \tag{B.5}$$

We now have a faithful representation of the *entire* Lie algebra of classical observables as formally self-adjoint operators on $\mathcal{H}_{pre}$.

Quantization

The pre-quantum Hilbert space is "too large". For the case in which Γ is a cotangent bundle, we are accustomed to representing quantum states by functions on the configuration space. Our pre-quantum Hilbert space consists, loosely speaking, of functions on the *phase space*; they are functions of too many variables. In the standard situation, only those elements of $\mathcal{H}_{pre}$ which are independent of momenta are to represent quantum states. Thus, we need a prescription by which we may restrict our attention to those pre-quantum states which depend on only half of the phase space coordinates. In general, this must be done locally, since elements of $\mathcal{H}_{pre}$ are not quite complex-valued functions. This is accomplished by choosing a *polarization* of maximal dimension. A polarization P is an assignment of a maximal subspace P_p of the complexified tangent space $T_p^{\mathbb{C}}\Gamma$ at each point that is:

i. Lagrangian: $\Omega_p(V,W) = 0 \quad \forall\, V, W \in P|_p$.

ii. integrable: If V^α and W^α are vector fields on Γ that belong to the polarization at each point, then $[V,W]^\alpha|_p \in P|_p \quad \forall\, p \in \mathcal{P}$;

If V is a vector field which, at each point, is an element of the polarization, we write $V \in P$.

The familiar polarization, which consists of the vertical tangent vectors of a cotangent bundle, is everywhere real. This is not the generic situation, however. Another useful example is a particular *complex* polarization, which is defined as follows. Suppose the phase space admits a Riemannian metric $g_{\alpha\beta}$ for which $J^\alpha{}_\beta := \Omega_{\beta\gamma}g^{\gamma\alpha}$ is an *almost complex structure*; i.e., we suppose that $J^\alpha{}_\gamma J^\gamma{}_\beta = -g^\alpha{}_\beta$, but do not require that J be integrable. The triple (Ω, g, J), if compatible in the assumed manner, equips Γ with what is called an *almost Kähler structure*. Since $J^2 = -\mathbf{1}$, its eigenvalues are $\pm$i. Now let P_p consist of all vectors $V \in T_p^{\mathbb{C}}\Gamma$ for which $J^\alpha{}_\beta V^\beta = \mathrm{i}V^\alpha$. Note that for any vector $w \in T_p\Gamma$, $W := w^\alpha - \mathrm{i}J^\alpha{}_\beta w^\beta \in P_p$, and that any element of P_p is of this form. It is easy to see that P satisfies condition (i) above. If, in addition, P is integrable, it is called a *Kähler polarization*. One can easily verify that if (Ω, g, J) equips Γ with the structure of a Kähler manifold, then P *is* integrable.

Having chosen a polarization P, one defines the *quantum wave functions* as the cross sections which are preserved by the polarization; $\xi : \Gamma \to E$ is a quantum wave function if it satisfies the *polarization condition*:

$$V^\alpha \mathcal{D}_a \xi = 0 \quad \forall\, V \in P.$$

The set $\mathcal{H}$ of such sections is to provide the 'physical' Hilbert space of the quantum theory. Unfortunately, the inner-product on $\mathcal{H}_{pre}$, whose only purpose was to render the operators $\hat{O}_f$ symmetric, does not, in general, induce one on $\mathcal{H}$. Construction of a *Hilbert space* from $\mathcal{H}$ is generically a non-trivial task. A lovely exception is the case in which P is Kähler. In this event $\langle , \rangle_{pre}$ directly carries over to $\mathcal{H}$.

For a given classical observable f, the pre-quantum operator $\hat{O}_f$ is, in general, not a well-defined operator on $\mathcal{H}$. If, however, $\hat{O}_f$ *does* preserve $\mathcal{H}$, f is called a *good observable*. The condition that

f be a good observable reduces to the statement that its Hamiltonian vector field *preserves the polarization*:

$$f \text{ is good} \quad \Leftrightarrow \quad \pounds_{X_f} V^\alpha \in P \quad \forall\, V \in P. \tag{B.6}$$

Note that, by use of the Jacobi identity, the Poisson bracket of two good observables is also a good observable. As a special case, if P is a Kähler polarization, the condition on the good observables is that, for any vector field v,

$$\begin{aligned} P \ni & \pounds_{X_f}(v^\alpha - \mathrm{i} J^\alpha{}_\beta v^\beta) \\ = & \pounds_{X_f} v^\alpha - \mathrm{i} J^\alpha{}_\beta \pounds_{X_f} v^\beta - \mathrm{i} v^\beta \pounds_{X_f} J^\alpha{}_\beta. \end{aligned} \tag{B.7}$$

Since the first two terms define an element of the polarization, any function whose Hamiltonian vector field preserves the complex structure is a good function.

References

[1] J.R. Klauder and B.S. Skagerstam, *Coherent states, applications in physics and mathematical physics*, (World Scientific, Singapore, 1985).

[2] A.M. Perelomov, *Commun. Math. Phys.* **26**, 222 (1972).

[3] R. Gilmore, *Ann. Phys. (N.Y.)* **74**, 391 (1972).

[4] R. Penrose, *The Emperor's New Mind*, (Oxford University Press, 1989).

[5] J.S. Bell. *Speakable and unspeakable in quantum mechanics*, p. 201, (Cambridge University Press, 1987).

[6] S. Weinberg, "Testing quantum mechanics", *Ann. Phys. (N.Y.)* **194**, 336–386 (1989).

[7] I. Bialynicki-Birula and J. Mycielski, "Nonlinear wave mechanics", *Ann. Phys. (N.Y.)* **100**, 62–93 (1976).

[8] A.M. Perelomov, *Generalized coherent states and their applications*, (Springer-Verlag, Berlin, New York, 1986).

[9] T.W.B. Kibble, "Geometrization of Quantum Mechanics", *Commun. Math. Phys.* **65**, 189–201 (1979).

[10] P.R. Chernoff, J.E. Marsden, *Properties of infinite dimensional Hamiltonian systems*, (Springer-Verlag, Berlin, Heidelberg, 1974).

[11] A. Heslot, "Quantum mechanics as a classical theory", *Phys, Rev.* **D31**, 1341–1348 (1985).

[12] J. Anandan and Y. Aharonov, "Geometry of quantum evolution", *Phys. Rev. Lett.* **65**, 1697–1700 (1990).

[13] G.W. Gibbons, "Typical states and density matrices", *Jour. Geom. Phys.* **8**, 147–162 (1992).

[14] L.P. Hughston, "Geometric aspects of quantum mechanics", in *Twistor Theory*, Ed. S.A. Huggett, (Marcel Dekker, 1995).

[15] R. Cirelli, A. Manià and L. Pizzocchero, "Quantum mechanics as an infinite-dimensional Hamiltonian system with uncertainty structure: Part I", *J. Math. Phys.* **31**, 2891–2897 (1990).

[16] R. Cirelli, A. Manià and L. Pizzocchero, "Quantum mechanics as an infinite-dimensional Hamiltonian system with uncertainty structure: Part II", *J. Math. Phys.* **31**, 2898–2903 (1990).

[17] C. Lanczos, *The variational principles of mechanics*, (Dover, 1970).

[18] V.I. Arnol'd, *Mathematical methods of classical Mechanics*, Second Edition, (Springer-Verlag, New York, 1989).

[19] Y. Choquet-Bruhat, C. DeWitt-Morette, M. Dillard-Bleik, *Analysis, manifolds and physics*, Revised Edition, (Elsevier Science Publishers B.V., 1987).

[20] J. Kijowski and W.M. Tulczyjew, *A Symplectic framework for field theories*, (Springer-Verlag, Berlin, Heidelberg, 1979).

[21] R. Geroch, *Mathematical physics*, (The University of Chicago Press, 1985).

[22] M. Reed and B. Simon, *Methods of modern mathematical physics*, v.I, *Functional analysis*, (Academic Press, 1980).

[23] P.A.M. Dirac, *The principles of quantum mechanics*, Fourth Edition, (Oxford University Press, 1958).

[24] P. Jordan, J. von Neumann and E. Wigner, "On an algebraic generalization of the quantum mechanical formalism", *Ann. of Math.* **36**, 29 (1934).

[25] A. Ashtekar, *Lectures on Non-perturbative canonical gravity*, (World Scientific, Singapore, 1991).

[26] N. Landsman, "Classical and quantum representation theory", hep-th/9411172 (1994).

[27] R. Shankar, *Principles of quantum mechanics*, (Plenum Press, New York, 1980).

[28] R. Penrose and W. Rindler, *Spinors and space-time*, Vol. 1: *Two-Spinor Calculus and Relativistic Fields*, (Cambridge University Press, 1984).

[29] A. Ashtekar, G.T. Horowitz and A. Magnon-Ashtekar, "A generalization of tensor calculus and its applications to physics", *Gen. Rel. Grav.* **14**, 411–428 (1982).

[30] M. Nakahara, *Geometry, topology and physics*, (Adam Hilger, 1990).

[31] S. Lang, *Introduction to differentiable manifolds*, (J. Wiley & Sons, 1962).

[32] R. Abraham, J.E. Marsden and T. Ratiu, *Manifolds, tensor analysis, and applications*, Second Edition, (Springer-Verlag, New York, 1988).

[33] P.A.M. Dirac, *Lectures on quantum mechanics*, (Yeshiva University, New York, 1964).

[34] M. Henneaux and C. Teitelboim, *Quantization of gauge systems*, (Princeton University Press, 1992).

[35] M.V. Berry, "Quantal phase factors accompanying adiabatic changes", *Proc. R. Soc. London* **A392**, 45–57 (1984).

[36] Don N. Page, "Geometrical description of Berry's phase", *Phys. Rev.* **A36**, 3479–81 (1987).

[37] R. Geroch, "A method for generating solutions of Einstein's equations", *J. Math. Phys.* **12**, 918–924 (1971).

[38] R.M. Wald, *General relativity*, (The University of Chicago Press, Chicago, 1984).

[39] K. Yano, *Structures on manifolds*, (World Scientific, Singapore, Philadelphia, 1984).

[40] L.P. Hughston, "Geometry of stochastic state vector reduction", preprint, Merrill Lynch and Kings College London, to appear in *Proc. Roy. Soc. London* (1995).

[41] J. Śniatycki, *Geometric Quantization and Quantum Mechanics*, (Springer-Verlag, New York, 1980).

[42] W. Rudin, *Real and complex analysis*, (McGraw-Hill, 1987).

[43] N. Woodhouse, *Geometric quantization*, (Oxford University Press, New York, 1980).

[44] R.M. Wald, "Quantum field theory in curved spacetime", in *Poceedings of Les Houches* (1992).

[45] A. Ashtekar and A. Magnon-Ashtekar, "A technique for analyzing the structure of isometries", *J. Math. Phys.* **19**, 1567–1572 (1978).

[46] M. Czachor, "Nonlinear Schrödinger equation and two-level atoms", *Phys. Rev.* **A**, 1–6 (1995).

[47] A. Messiah, *Quantum mechanics*, (North-Holland Amsterdam, 1961).

[48] W.-M. Zhang, "Coherent states: Theory and some applications", *Rev. Mod. Phys.* **62**, 867–927 (1990).

[49] J.R. Klauder, *J. Math. Phys.* **4**, 1055 (1963), *J. Math. Phys.* **4**, 1058 (1963).

[50] H. Goldstein, *Classical Mechanics*, (Addison-Wesley, 1981).

[51] J. Lee, "Geometric Quantization", in *New perspectives in canonical gravity*, Ed. A. Ashtekar, (Bibliopolis, Napoli, 1988).