COMPLEX NUMBERS IN QUANTUM THEORY

Glenn Maynard

Dissertation Prepared for the Degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

August 2015

APPROVED:

William D. Deering, Major Professor
Duncan Weathers, Committee Member
Paolo Grigolini, Committee Member
Yuri Rostovtsev, Committee Member
Carlos Odonez, Interim Chair of the Department of Physics
Mark Wardell, Dean of the Toulouse Graduate School

Maynard, Glenn. Complex numbers in quantum theory. Doctor of Philosophy (Physics), August 2015, 38 pp., 2 figures, 12 numbered references.

In 1927, Nobel prize winning physicist, E. Schrodinger, in correspondence with Ehrenfest, wrote the following about the new theory: "What is unpleasant here, and indeed directly to be objected to, is the use of complex numbers. Psi is surely fundamentally a real function." This seemingly simple issue remains unexplained almost ninety years later.

In this dissertation I elucidate the physical and theoretical origins of the complex requirement. I identify a freedom/constraint situation encountered by vectors when, employed in accordance with adopted quantum representational methodology, and representing angular momentum states in particular. Complex vectors, quite simply, provide more available adjustable variables than do real vectors. The additional variables relax the constraint situation allowing the theory's representational program to carry through.

This complex number issue, which lies at the deepest foundations of the theory, has implications for important issues located higher in the theory. For example, any unification of the classical and quantum accounts of the settled order of nature, will rest squarely on our ability to account for the introduction of the imaginary unit.

Copyright 2015

By

Glenn Maynard

ACKNOWLEDGEMENTS

In thanks to and memory of

Robert Weingard, PhD

and

Gertrude E. Row.

TABLE OF CONTENTS

Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In this dissertation I present an explanation of the requirement for complex numbers in the mathematical formulation of quantum mechanics.

This problem is located deep in the foundations of quantum theory and historically, it is raised by the founding fathers within "the first moments" after the birth of the theory.

There are a number of ways one can formulate the complex numbers issue. Complex numbers and the imaginary unit are encountered throughout the mathematical theory. One can choose any occurrence and ask there for an explanation. Why does the Schrodinger equation of motion prominently display the imaginary unit? Why is an "I" encountered in the operators associated with physical observables? Why is "I" encountered in the operator commutation relations, in the components of the state vector, etc.?

The problem can be seen, however, in a quite stark way by formulating it in full generality as a point of logic. Consider the following list of four propositions. Each seems obviously true, yet the four generate a contradiction.

A Point of Logic:
1. Physical theories describe physical objects, object interactions, and evolutions.
2. All description of physical objects, object interactions, and evolutions, requires real numbers only.
3. It follows from 1 and 2 that physical theories require real numbers only.
4. Quantum theory, however, requires complex numbers.

Presumably, if one starts from a clean slate and undertakes to logically reconstruct the quantum mathematical formal structure, then there will come some specific point and some specific reason for which complex numbers must be introduced into the theory. In this dissertation, I will identify that point and make clear the origin of the complex requirement.

It will be a central point of the dissertation that quantum theory adopts a methodology of using vectors (mathematical objects) to represent object states (physical phenomena) that is different from the way that vectors are used in previous physical theories. I will show that it is this method of using vectors that has a traceable connection to the requirement that the vectors be complex.

The theory's adopted representational conventions, adopted by postulate, provide a useful separation line for the following explanation. The explanation I give in chapter 3 begins with the postulates and demonstrates the requirement for complex state vectors. Chapter 4 considers the other side of the separation line, that is, the origins of the postulates. It is useful explanatory headway to see the complex requirement demonstrated from the postulates (and certain features of the observed phenomena). However, additional explanatory insight is provided if we also have some understanding of the physical and representational principles which underlie and motivate the postulates. Chapters 3 and 4 address both sides of the separation line.

Before the business of explaining, chapter 2 collects together some of the reasons why this issue, the requirement for complex numbers in quantum theory, is an important issue. Chapter 2 introduces other research both contemporary and historical which either addresses directly or is related to the complex number requirement.

Please enjoy.

CHAPTER 2

IMPORTANCE AND RELATED ISSUES

2.1 Importance of the Complex Requirement

In this section, I discuss the importance of the issue of complex numbers in quantum theory. The issue is, of course, important in its own right. In addition, however, it has tight connections and implications for important secondary issues. These points are discussed below. I also introduce other related research both contemporary and historical acknowledging the ongoing importance of this issue. In the following I separate out and discuss individually several of these points.

2.1.1 Familiarity versus Understanding

In the introduction, I listed four propositions which starkly present the complex number issue as a point of logic. The four propositions seem obviously true, yet, nonetheless, are obviously inconsistent. It is therefore a remarkable fact that at this moment in history, there is little awareness of or interest in this issue. Only a very small contingent would consider the issue important.

This, I think, is because of the particular history of the issue. The apparent requirement for complex numbers attracted the attention of the founding fathers. Without quick resolution, however, it was put aside and physicists went about the business of actually using the theory. Now generations of physicists have begun and ended professional careers with the issue well buried in the background.

What has happened is that physicists have grown accustomed to the presence of complex numbers in the theory. Given this comfort, inquiry from a student can be interpreted as a student's discomfort arising from lack of familiarity with complex numbers rather that as a legitimate question about the theory. Consequently, even opening the question for discussion is, in effect, discouraged. This situation is particularly puzzling because in the context of other theories, it is considered important that the student understand the role of any complex numbers present. Typically that role is to provide a convenience in mathematical manipulations. The situation is different, however, in introducing quantum theory. The student takes the lesson that apparently some things are opaque and no explanation will be forthcoming. Fundamental features of the theory become essentially established theory dogma. This process does not serve the student well in shaping an intellectual and scientific mindset.

It is essential therefore that we recognize the following distinction. To grow accustomed to something is different from understanding that thing. That a physicist is comfortable with the complex numbers or has grown accustomed to them is of no concern. None of physics is about the psychological state of the physicist or her comfort level with this or that feature of the theory. The issue is whether what the theory says is, in fact, understood. In this case, the significance of the "i" in the theory is not understood.

That the psychological considerations blind us to our failure to recognize the complex question can be seen clearly by considering the situation confronting Schrodinger. He was clearly an intelligent, thoughtful, physicist who certainly had some good insight into quantum physics. For him, however, the complex requirement was new. He had not "grown accustomed". It is important, therefore, to weigh heavily this Nobel prize winning physicist's analysis of the apparent requirement for complex numbers in the theory. Clearly he recognized an issue that required explanation.

> About (a), Schrodinger said that he had abandoned the expression $\Psi(\partial\Psi^*/\partial t)$ of his earlier manuscript (Schrodinger, 1926f), and was now focusing on $\Psi\Psi^*$ for the electric charge density in real space. He then continued: "What is unpleasant here, and indeed directly to be objected to, is the use of complex numbers. Ψ is surely fundamentally a real function." There followed an involved suggestion of how to generate a complex Ψ from its real part Ψ_r, a suggestion clearly not quite satisfactory to Schrodinger himself.[1]

The history of science provides us with stark examples of the conflation of familiarity with understanding. There is an important lesson to be learned from those who took the earth to be flat yet observed the sun's daily passage. To be accustomed to something is not to understand that thing.

2.1.2 The Value of Paradox

The obvious importance of the complex numbers in the theory is that, presumably, they are not there gratuitously. That is, they are there for a reason. They are therefore a large red flag indicating that here is an open question that requires explanation. The theory is saying something. What is it saying?

Professor Yakir Aharonov emphasizes the value of these red flags. He makes the point in terms of the value of paradox in physics. By paradox, he intends, for example, the set of four apparently true but inconsistent propositions which I listed in the introduction.

> We will use paradox to probe quantum physics. Can paradox be useful? The history of physics shows how useful. As Wheeler put it, "No progress without a paradox!"[2] …
>
> A paradox is an argument that starts with apparently acceptable assumptions and leads by apparently valid deductions to an apparent contradiction. Since logic admits no

> contradictions, either the apparently acceptable assumptions are not acceptable, or the apparently valid deductions are not valid, or the apparent contradiction is not a contradiction. A paradox is useful because it can show that something is wrong even when everything appears to be right. It does not show what is wrong. But something is wrong – something we thought we understood – and a paradox moves us to reexamine the argument until we find out what is wrong.[3]

2.1.3 An Issue Located in the Foundations

In this dissertation, I explain that the complex number requirement is introduced in construction of the state vector. This means that if we do not understand the complex requirement, then essentially we do not understand the quantum state. The fact that the quantum state is not fully understood is not, in fact, news. It is the accepted status quo. Until students are trained to not ask, they do wonder about the “cloud” around the atom, and whether the wave function is something like an E or B field only somehow complex valued.

To fail to understand something about the quantum state is to introduce misunderstanding at the deepest levels of the theory. Reason dictates that incomplete or incorrect understanding at this level will have consequences throughout the theory.

First consider the importance of the complex number issue for understanding the state. This offers a perfect example of the usefulness of the complex number issue. In the following chapters, I explain that the complex number issue is, in fact, a secondary issue located in the context of the larger issue of how the theory employs vectors to represent states of objects. That is, the theory first makes a commitment to the use of vectors to represent states. It is then a secondary consideration why those vectors must be either real or complex. It is the complex number issue that draws our attention to the more general issue that the quantum concept of state is importantly different from the classical state. I discuss the particular quantum concept of state in chapter 4.

As mentioned, the implications of the complex number issue are not limited to insights at the level of the state. As we go higher in the theory from the state to the equation giving the motion of the state, we again see the importance of the complex number requirement. It has been an important program since the birth of the theory to give some account of the connection between quantum and classical theory. In particular, there have been a variety of attempts to explain the Schrodinger equation of motion in terms of classical equations of motion and vice versa. It is a point of logic, however, that any explanation of one in terms of the other must include some account of why the imaginary unit is introduced in the Schrodinger equation. That is, this unification can not go forward without some explanation for the requirement for complex numbers in quantum theory. Presumably, the unification can not successfully go forward without a correct explanation.

2.2 Related Issues

To provide some broader sense of the importance of the complex requirement, I briefly mention some questions and issues that are intimately linked.

First: Quantum theory adopts a new way of using vectors to represent physical phenomena that is different from the use of vectors in previous theories. Failure to recognize this fundamentally different representational use of vectors constitutes a confusion about the state vector and the representational methodology employed by the theory. As I discuss, the requirement for complex vectors is closely associated with and a consequence of this new methodology.

Second: Are the complex numbers present in the theory because they reflect some complex feature of the observed physical phenomena? If so, what does this mean for the observed phenomena to be complex? Are they present instead to provide some mathematical convenience? If so, what is the specific nature of the convenience? Do they simplify mathematical manipulation of oscillatory expressions as they do in classical theories?

Third: Are complex numbers somehow required in expressing the dynamics of the physical phenomena? I discuss that this is an explanation avenue that has been considered by a variety of researchers. Whether the complex requirement is associated with dynamics or prior to dynamics can have sweeping implications for this long and distinguished research program.

2.3 Related Research

In this section, I briefly review two important lines of research which rest squarely on the complex number issue.

2.3.1 Logical Reconstruction of the Quantum Mathematical Formal Structure

There is an ongoing line of research that undertakes to provide a full explanation of all features of the quantum mathematical formalism. The driving methodology is to elucidate the physical origins of all structural details of the mathematical theory. One part of this program, of course, would be to explain the requirement of the complex numbers in the theory. This particular issue is addressed by Goyal, Knuth, and Skilling in an article entitled, “Origin of complex quantum amplitudes and Feynman's rules".[4] The article abstract describes their program.

> Complex numbers are an intrinsic part of the mathematical formalism of quantum theory, and are perhaps its most characteristic feature. In this paper, we show that the complex nature of the quantum formalism can be derived directly from the assumption that a pair of real numbers is associated with each sequence of measurement outcomes, with the probability of this sequence being a real-valued function of this number pair. By making use of elementary symmetry conditions, and without assuming that these real number pairs have any other algebraic structure, we show that these pairs must be manipulated according to the rules of complex arithmetic. We demonstrate that these complex numbers combine according to Feynman's sum and product rules, with the modulus-squared yielding the probability of a sequence of outcomes.[4]

In this article, the authors introduce by postulate the real number pairs and then propose an explanation why the pairs obey the complex number combination rules. They state that the explanation relies not just on "elementary symmetry conditions" but also on physical assumptions.

> Using symmetry and consistency conditions that arise naturally in an operational framework, and making a few elementary physical assumptions, we show that this postulate leads to Feynman's rules of quantum theory. Most importantly, we show that the number pairs assigned to each sequence of measurements outcomes must be manipulated according to the rules of complex arithmetic, without assuming this at the outset.[4]

How does this work relate to the complex number explanation that I provide in this dissertation? My explanation takes as a starting point that the theory adopts a particular methodology for using vectors to represent a certain class of physical phenomena. I explain that an essential requirement for use of this methodology is the use of vector spaces with a defined L^2 vector norm. It is this commitment to the L^2 vector norm that explains why, if the vectors have two real variables per component, then they must obey the complex number combination rule. On my explanation, the combination rule is a consequence of the representational choice to use an L^2 vector space. Goyal, however, explains the combination rule using "symmetry and consistency conditions … and a few elementary physical assumptions."[4]

Consequently Goyal and I offer fundamentally different explanations for the complex number requirement. It would be an important topic for a follow on analysis to determine whether Goyal's symmetries, consistency conditions, and physical assumptions, in fact, also import a requirement that the vector space be defined as L^2. That is, does Goyal's explanation assume the quantum representational methodology without explicitly acknowledging doing so?

2.3.2 Unification of Classical and Quantum Equations of Motion

As mentioned, unification of quantum and classical theory has been an ongoing historically, and conceptually important program. A main focus of the program is to interchangeably derive the classical from the quantum equations of motion. A recent article, by Schleich, Greenberger, Kobe, and Scully revisits this program.[5]

In the course of relating the two theories' equations of motion, Schleich et. al. introduce the imaginary unit as follows:

> In the past, Stückelberg[6], Wheeler[7], and many others have addressed the question of why the imaginary unit appears so prominently in quantum mechanics but not in standard formulations of classical mechanics. However, the complex-valued function ψ (cl), which obeys the nonlinear Schrödinger equation, demonstrates that the appearance of the imaginary unit i is not a characteristic feature of quantum mechanics but, rather, reflects the fact that the underlying dynamics rest on two equations rather than one: the continuity equation and the Hamilton– Jacobi equation. At this point, it is of no importance that the latter implies the former. Therefore, complex numbers are just a useful tool to combine two real equations into a single complex equation. We conclude our discussion of classical matter waves by noting that the appearance of i in quantum mechanics as a mathematical convenience rather than a necessity is also

confirmed by the formulation in terms of the Wigner phase space distribution function[8]. Indeed, this quantity is always real. We return to this point in Conclusions and Outlook.[5]

How does this account of the origin of the complex requirement relate to the one that I give in the dissertation? On the Schleich account, the dynamics of the system "rest on two equations rather than one". It is therefore an issue of the dynamics of the system that is the origin of the imaginary unit requirement. In the explanation that I give, the requirement for the imaginary unit arises prior to any consideration of dynamics. Instead, the complex numbers are introduced in the logical construction of the state vector (for a particular moment in time) for the purpose of relaxing a constraint encountered in vector pairwise relationships.

In this program, ie., presenting a unified understanding of the equations of motion, the correct origins of the complex requirement play a crucial role. If the complex requirement is, in fact, prior to dynamics, then we can say, in one stroke, that any introduction of the complex requirement as due to dynamics would appear to be incorrect. If the state is complex prior to dynamics, then the explanatory connection must go the other way. It is the complex state that explains the combining of the two equations, not the combining of the equations that explains the complex state.

This issue is much larger than just the Scheich article. The Schleich article is one link in a very traceable chain employing the same mathematical machinery although sometimes to make different points. The same mathematical arguments can be observed in Bohm's[9][10] early hidden variable program and in Holland's[11] later revisit of that program.

With respect to this entire historically important line of research, much rests on whether the complex requirement is associated with system dynamics or whether, in fact, it is prior to dynamics as I propose.

CHAPTER 3

THE COMPLEX NUMBER EXPLANATION

3.1 Preliminary Comments

This chapter presents the explanation of the requirement for complex vectors in the quantum mathematical formal structure. The explanation takes as starting points, 1., the theory's adopted conventions for using vectors to represent observed state transition probabilities, and, 2., empirically observed features which characterize angular momentum states of physical objects. The chapter begins with some preliminary comments regarding the explanation.

3.1.1 The Principle of Unique Causes

The focus of this chapter is on the vectors used by the theory to represent states of physical objects. The dissertation task, however, is to explain the presence of complex numbers not just in association with a particular theoretical feature, but in full generality throughout the theory. How can this task be fulfilled if attention is restricted to just state vectors?

I take the view, seemingly demanded by reason, that there is some single explanation requiring the introduction of complex numbers into the theory. All other occurrences throughout the theory are then simply mathematical consequences. This "principle of unique causes" seems a point of general principle and governing across physics. As an example, suppose the reason a planet orbits a sun is due to a gravitational relationship. Does reason allow a second genuinely logically distinct reason for the orbit? If a second were proposed, we would expect that they were not genuinely logically distinct but only seemingly so and, in fact, explainable one in terms of the other.

In a logical reconstruction of quantum theory, presumably there is some specific point at which, and reason for which the theorist must introduce complex numbers. The position that I take here is to say that once required, there is no second, additional, genuinely logically distinct reason that would again require their introduction. A corollary to this would be that given two seemingly different reasons for complex numbers in the theory, the two are in fact the same and each can be explained in terms of the other.

In the remainder of the chapter I explain that the point at which complex numbers are introduced into the theory is in the logical construction of the state vector. It is at this point that we can identify reasons external to the theory that require their introduction. Occurrences of complex numbers in other mathematical structures associated with the theory are then simply mathematical consequences of the state vector being a complex vector. That is, the imaginary unit is present in the equation of motion, the operators, the commutation expressions, etc., only as a mathematical consequence of the state vector being a complex vector.

I therefore continue with full attention turned to the state vector.

As a second preliminary comment, I mention to the reader that the explanation is presented in two parts. The dividing line that separates the two is the set of conventions adopted by the theory for the mathematical representation of physical phenomena.

What exactly do I mean by the theory's representation conventions? Any mathematical physical theory undertakes to represent physical phenomena by using mathematical structures and objects. To do this, the theory must adopt some convention for the representation, that is, exactly how it will mathematically represent physical phenomena. Since a convention is just that, a convention, any such choices made by the theorist must be introduced or adopted by postulate. We therefore recognize two classes of postulate in theory construction. First, a postulate must be present to adopt any chosen representational convention. The quantum convention to represent object states by using Hilbert space vectors is adopted by such a postulate. Alternatively, we see that the postulate that the state evolve in accordance with the Schrodinger equation is of a different class. This postulate states a matter of physical fact. In theory construction, it is the theorist's choice how to mathematically represent the state. Having done so, however, how the state evolves is a matter of empirical fact.

In order to identify the representational conventions adopted by quantum theory, we, therefore, turn to the postulates. Quantum theory makes a very early and fundamental commitment to, 1., the use of vectors to represent states, and, 2., a "Born Rule" principle to represent state transition probabilities.

As stated above, I separate the complex vector explanation into two parts. Using the theory's adopted representational conventions, ie., the representational postulates, as the separation line, I separate the explanation as follows.

3.1.2 This Chapter – The Requirement for Complex Vectors

In this chapter, the explanation runs from the conventions to the complex vectors. That is, the explanation takes as the starting point that the theory adopts the representational conventions just mentioned. From this starting point, (and a secondary consideration, ie., the requirement that the representational vectors vary as the actual physical states vary), I explain why the vectors need to be complex.

3.1.3 Next Chapter – The New Quantum Representational Methodology

One could say that the above explanation is open to a somewhat valid criticism. That is, the explanation is anchored only in the theory postulates. With respect to the task of "explaining", that is, "answering the why question", the criticism is, the postulates are left as impenetrable and unexplained. In one sense, this is as far as one can go in explaining. To say that a method of representation is adopted by convention, ie., by postulate, means that it is fundamentally nothing more than a choice made by the theorist during theory construction. A different theorist might have made different choices. No where do we see the adoption of alternative conventions by different theorists more clearly demonstrated than in the historical development of quantum theory. The explanation, therefore, could respectably end by simply saying "theorist's choice".

On the other hand, conventions, presumably, are not adopted whimsically. It seems reasonable that a theorist would choose a representational convention that is in some way well suited to the features of the physical phenomena to be represented.

If our program is to explain the quantum requirement for complex numbers, we certainly have the right to anchor the explanation in the conventions. If however, some insight is possible regarding why the particular conventions are adopted, then a broader understanding of the complex requirement may be available.

For this reason, the next chapter looks "behind the curtain" and discusses the considerations which motivate the adoption of the quantum conventions. The dissertation is meant to be explanatory, that is, provide insight, understanding, and some intuitive feel for the complex requirement. That goal is furthered by not simply recognizing that the theory adopts particular conventions, but also shining light on why these conventions are adopted.

This chapter, together with the "look behind the postulates curtain" provided in the next chapter, fully explain the quantum requirement for complex vectors.

3.2 The Requirement for Complex Vectors[a] [12]

3.2.1 Section Intro

In this section, as just discussed, I take as the starting point that the theory adopts, by postulate, certain conventions for the use of mathematical objects for the representation of physical phenomena. That is, the theory is committed to the use of vectors to represent states, and the Born rule to represent state transition probabilities. From this starting point, I then explain why the vectors representing the states must be complex vectors.

I have constructed the explanation around a set of four requirement statements, R1-R4. These are meant to provide a solid vantage point on which the reader can stand and survey the explanation logic. This vantage point provides two perspectives meant to be considered independently.

Perspective 1 – Satisfying the Requirements: From one perspective, the four requirements can be considered simply a set of abstract requirements on the vectors of a

[a] Material from section 3.2 is reproduced in part from, Maynard, G., Lambert, D., Deering, W. D., (2015), (ref. endnote 12), with permission from Physics Education (IAPT).

vector space. Taking this perspective one can follow, as a purely mathematical issue of vector structure, that the set of requirements is not satisfied by real vectors, but can be satisfied if the vectors used are complex.

Perspective 2 – Origins of the Requirements: Taking the second perspective, I put aside considerations of mathematical consequences and instead trace the origins of the four requirements. I show that two of the requirements, R1 and R2, originate in the conventions adopted by the theoretician during theory construction. The remaining two requirements, R3 and R4, originate from physical considerations, in particular, the requirement that a vector employed to represent a physical state must have the ability to vary as the physical state itself varies.

To begin, I list the set of four requirements, R1 – R4. I then discuss, in turn, first satisfying the requirements, and then the origins of the requirements.

3.2.2 The Set of Requirements: R1-R4

Let V be an arbitrary vector in a finite dimensional vector space. R1 – R4 are requirements on vector V.

R1: Vector, V, is subject to n independent constraints with respect to an orthonormal basis set of vectors $\{ b_i \}_{(i = 1, n)}$. Each constraint is of the form, $p_i = | (V , b_i) |^2$ (where the parenthesis indicates inner product).
R2: Vector V is n-dimensional.
R3: Vector V must vary with two real variables, r and c.
R4: The set of constraints mentioned in R1 vary parametrically with the variable r above. That is, $p_i = p_i (r)$

3.2.3 Satisfying the Requirements

Here I consider the four requirements simply as a set of abstract requirements on the vectors of a vector space. I want to show, as a mathematical issue, that the set of requirements is not satisfied by real vectors, however, may be satisfied by complex vectors.

I show first that real vectors do not satisfy these requirements:

Point 1 - Number of variables present: Assume vector V is a real vector. Real vectors vary with one real variable per vector dimension. Requirement R2 requires that V is n-dimensional. Therefore, V varies with n variables.

Point 2 - Number of constraints present: We see from R4 that the set of R1 constraints vary parametrically with variable r. If we consider any fixed value of r, then the R1 constraints impose n independent constraints on V (with respect to the orthonormal basis set { b_i }).

Points 1 and 2 imply: It follows that vector V is fully specified with respect to the basis set { b_i } (for any fixed r). There are n variables and n independent constraints. All variables present are assigned values by the constraints.

Therefore: If vector V is real, and satisfies requirements, R1, R2, and R4, then it:

a. Is fully specified by r, and

b. Varies as a function of r.

Therefore: Having satisfied requirements, R1, R2, and R4, vector V cannot satisfy requirement R3. Vector V varies with and is fully specified by r. Consequently, it is not possible for the vector to vary (nontrivially) in the second variable, c, as is required by requirement R3.

I have shown that if vector V is real then it does not satisfy the set of requirements. Having done this analysis, however, one sees how substituting complex vectors for real vectors avoids the constraint limitation. The constraint encountered by real vectors is due to the availability of only n variables in the face of n constraints. A complex vector, however, provides 2n independent real variables. In the face of only n constraints, a 2n variable complex vector provides sufficient freedom to vary in both r and c degrees of freedom.

3.2.4 Origins of the Requirements

I next return to the set of four requirements and consider where the requirements originate. I show that the four requirements have two origins, two vector requirements from one origin and two from the other. Requirements, R1 and R2, are consequences of the theory's adoption of a particular representational convention. Their origin is therefore from particular choices made by the theoretician during theory construction. Requirements R3 and R4 are consequences of observed physical phenomena. Here I invoke the general principle that to mathematically represent a physical state, the mathematical object chosen for the representation must have the ability to vary as the actual physical state varies. In the present situation, the quantum state vectors must have the ability to vary with the freedom of angular momentum states.

I discuss the requirements R1-R4 in turn and identify their individual origins.

3.2.4.1 Origin of Requirement R1

Requirement R1: Vector, V, is subject to n independent constraints with respect to an orthonormal basis set of vectors $\{ b_i \}_{(i = 1 , n)}$. Each constraint is of the form, $p_i = |(V , b_i)|^2$.

In constructing the mathematical structure of any physical theory, some convention must be adopted for the representation of physical phenomena by mathematical structures. Quantum theory adopts, by postulate, the following representational convention:

P1: States of physical objects are represented by unit vectors, V.

P2: The probability for a transition between two states is represented by the "Born Rule". The "Born Rule" yields the probability as an inner product function on two vectors, V and b, which represent the two physical states involved in the transition, $p = |(V, b)|^2$.

The important point that we recognize in this section is that adoption of the Born Rule, in fact, imposes a constraint on vector V relative to vector b.

The familiar use of the Born Rule is to enter with the two state vectors, V and b, and obtain the transition probability, p. Here, we are recognizing a different perspective. It is the probability value that is the observed physical fact. The vectors are merely mathematical structures employed to represent physical states. By adopting the Born convention to represent transition probabilities we are required to choose vectors which yield the correct inner product value equal to the probability represented. From this perspective, the Born Rule, in fact, defines the vector pair (a partial definition) by specifying their inner product relationship. Consequently, we recognize a Born Rule expression as a "Born Constraint" on a state vector, V, relative to a transition state vector, b.

In addition to recognizing that the Born Rule imposes Born Constraints, requirement R1 also claims that there are n independent Born Constraints (with respect to the basis set). How do we know this?

A Born expression, $p = | (V , b) |^2$, represents the probability for a single transition from one state to another. It is observed physical fact, however, that an object in a given state can transition into one of some number, n, of alternative possible transition states. Each transition has some observed probability, p_i , and since they are mutually exclusive and exhaustive, $\Sigma_{(i = 1 , n)} (p_i) = 1$. This fact about probabilities imposes a requirement on the set of Born expressions representing the probabilities for the set of possible transitions.

That is:

$$1 = \Sigma_{(i = 1 , n)} (p_i) = \Sigma_{(i = 1 , n)} (| (V , b_i) |^2) , \quad \text{(Eqn. 1)}$$

We can recognize this as, in fact, a requirement on the vector space used to represent states by the Born rule. That is, the vector space must come equipped with a defined L^2 vector norm.

The L^2 vector norm is defined as follows:

$$| V | = \Sigma_{(i = 1 , n)} (| (V , b_i) |^2) .$$

If vector V is a unit vector, then,

$$1 = \Sigma_{(i = 1 , n)} (| (V , b_i) |^2). \qquad \text{(Eqn. 2)}$$

We see then that the choice to represent a set of transition probabilities by the Born Rule (Eqn. 1) has imposed the requirement that the vector space must be defined to have an L^2 vector norm (Eqn. 2).

Recognizing that the vector space has an L^2 vector norm is useful as follows. The set of vectors $\{ b_i \}_{(i = 1 , n)}$ in (Eqn. 2) are an orthonormal basis set. Consequently there is a set of n individual Born Constraints on vector V, one associated with each basis vector. These constraints are independent because each is relative to a basis vector that is orthogonal to all of the others.

We therefore have the result R1 stated above. Any vector V used by the theory to represent the state of an object must satisfy requirement R1.

I note, in passing, an interesting and pedagogically useful point. The explanation of why quantum theory employs Hilbert space vectors to represent states is sometimes opaque. Here we understand that the theory makes an, early and fundamental commitment to the use of vector spaces which have an L^2 structure. If one generalizes the structure of a vector space in every way, dimensionality, etc., but retains the L^2 structure, then that is the set of Hilbert spaces. Quantum theory employs Hilbert spaces because the theory makes use of and therefore requires the L^2 structure.

3.2.4.2 Origin of Requirement R2

Requirement R2: The vector is n dimensional.

Having done the work of the previous section, we immediately recognize this requirement on any state vector. As explained, vector V is in a vector space spanned by the n orthonormal basis vectors $\{ b_i \}_{(i = 1\ ,\ n)}$. Consequently, V is n dimensional.

3.2.4.3 Origin of Requirement R3

Requirement R3: Vector V must vary with two real variables, V (r, c).

It is observed physical fact that angular momentum states vary as a function of orientations or directions in physical space. The point is general, but can be seen by considering a simple example of two spin 1/2 objects. Suppose one object interacts with a

Stern-Gerlach apparatus oriented in the z direction and deflects up along that direction. The second object interacts with a machine oriented along the (θ , ϕ) direction and deflects up along that direction. Subsequent to the interactions, these two objects are in objectively different physical states. What does it mean to be in different states? It means that subsequent observations made on the objects will yield different results (observations are on ensembles). They are observably different. We can state this same physical fact in another way by saying that angular momentum states vary with orientations in physical space.

It is a general point that in constructing a mathematical theory, for any mathematical object chosen to represent the physical state, then that mathematical object must have the ability to vary as the physical state does. In particular, any vector we employ to represent angular momentum states must have the ability to vary with orientations in physical space. We can recognize this explicitly by writing the state vector as a function of orientation, V (O), where "O", see Figure 1, is an orientation in physical space.

Orientations in three dimensional physical space vary with two degrees of freedom. Typically, polar coordinates, (θ , ϕ), are chosen to label spatial orientations. Here it is useful to choose different coordinates as depicted in Figure 1. Select an arbitrary orientation, O_2 , then let real variables (r , c) label variation in radial and circumferential degrees of freedom relative to O_2 .

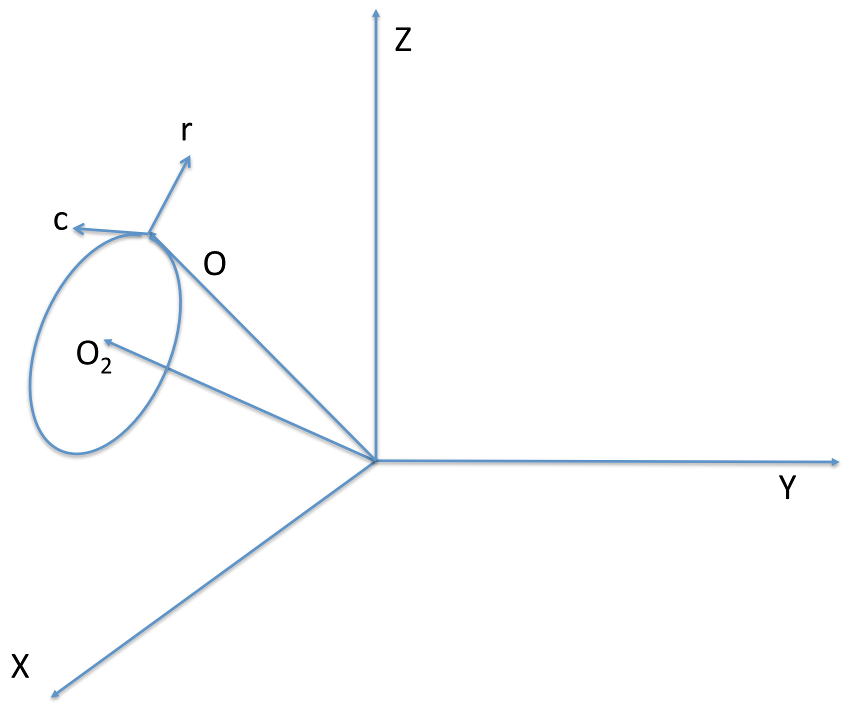

Figure 1. Physical space orientations O_2 and O. Defined coordinates r and c.
Coordinates r, (radial) and c (circumferential) are defined relative to O_2.

We can explicitly recognize this variation in two degrees of orientation freedom by writing the above state vector, V (O), as V (r , c) with r and c coordinates as defined.

We therefore have requirement R3 as given above. Note that the point here is to recognize that any vectors representing angular momentum states must have the ability to vary as the actual physical state varies, i.e., with two orientation degrees of freedom.

3.2.4.4 Origin of Requirement R4

Requirement R4: The set of constraints mentioned in R1 vary parametrically with the variable r above. That is, $p_i = p_i (r)$

(Since requirement R4 references the R1 Born Constraints, we copy again R1.

Requirement R1: Vector, V, is subject to n independent constraints with respect to an orthonormal basis set of vectors $\{b_i\}_{(i = 1 , n)}$. Each constraint is of the form, $p_i = |(V , b_i)|^2)$

In the last section, we recognized the physical fact that angular momentum states vary with physical space orientations. Here, we recognize a second empirically observed fact characterizing angular momentum states. That is, for two angular momentum states associated with two different physical space orientations, O_1 , and O_2 , the probability for a transition from one state to the other varies as a function of the separation angle between the two orientations.

Here is where we can take advantage of the r and c coordinates defined earlier. If we take O_2 to be our arbitrary fixed reference, then the separation angle between the two orientations, O_1 , and O_2 , is given by the coordinate r. Consequently, $p_i = p_i (r)$.

For the Born Constraints to vary parametrically with r we have made an assumption. That is, vector V is associated with one spatial orientation, O_1 , and all of the transition state vectors, $\{ b_i \}$ are associated with a single orientation, O_2 . This is appropriate for angular momentum observations. Suppose an object is in the state represented by vector, $V (O_1)$. The object then interacts with a Stern-Gerlach apparatus oriented along O_2 . In this case, there are a set of n possible transition states, but we note the important fact that they are all associated with physical space orientation, O_2 .

Consequently, we have the result that the initial state vector is subject to a set of Born Constraints relative to the transition state vectors, $\{ b_i \}$, and these constraints all vary parametrically with the separation angle parameter r.

CHAPTER 4

THE NEW QUANTUM REPRESENTATIONAL METHODOLOGY

4.1 Introductory Comments

The previous chapter provided an explanation of why the state vectors employed by the theory must be complex vectors. I used as a starting point for this explanation the representational postulates, ie., the fact that the theory adopts a particular method of using vectors to represent states and state transition probabilities.

In this chapter, I turn attention to those adopted conventions. It is a central point in the following discussion that quantum theory employs vectors to represent physical phenomena in a way that is different from previous physical theories. Here, I discuss how this new methodology works and why it is adopted by the theory.

What do we need to understand about the quantum adopted vector use methodology? Recall that the explanation of the last chapter rested on two stated requirements, R1 and R2, that trace their origins to the theory's adopted conventions. I emphasized that the main insight leading to R1 is that the Born rule convention employed by the theory also, in fact, constitutes a Born rule constraint on one vector with respect to another. It is this fact that the theory's methodology imposes vector pairwise constraints that I showed leads quite directly to the requirement for complex vectors when, in particular, the vectors are employed to represent angular momentum states. It is therefore of some explanatory use to understand the origins of the Born rule. What considerations would prompt a theorist, in theory construction, to adopt this representation convention?

Adoption of the Born rule is fundamentally the exploitation of the L^2 vector norm structure defined on a vector space for the purpose of representing probability distributions. That is, there is an isomorphism in structure that is available relating the two, ie., probability distributions and the L^2 vector norm, that can be employed as the basis for a representational technique. The principle can be shown in an intuitive way by using a simple example.

4.2 The Simple Intuitive Example

In order to get some intuition for the new quantum use of vectors, I begin with a simple example.

Consider a physical object which can undergo probabilistic transitions from some initial state into one of some set of alternative possible transition states. To construct a simple visualizable example, we can consider the slightly contrived situation where we toss a six-sided die, but disregard all tosses which yield other that one, two, or three. For an ensemble of tosses, there is some observed probability, p1, p2, and p3, associated with each of the possible alternatives, one two, and three.

We can mathematically represent this set of alternative possible transitions and the associated probabilities by using vectors in the following way.

First - Select an arbitrary orthonormal basis set of vectors in a three dimensional real vector space which is equipped with the familiar Euclidean, ie., L^2, vector norm.

Second - Define a pairwise association between the three possible transitions, ie., outcomes, one, two, and three, and the three basis vectors, B1, B2, and B3, as depicted in Figure 2 below. By doing this, we employ a vector to represent a possible toss outcome.

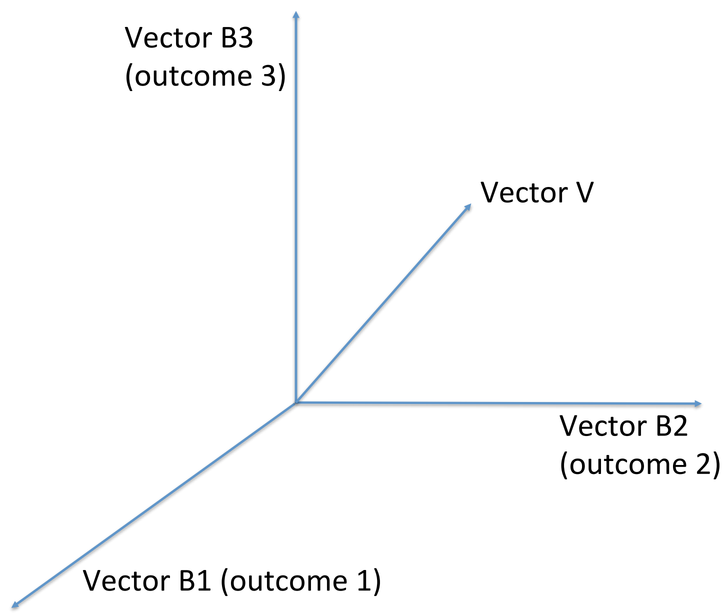

Figure 2. The quantum representational methodology

- Associate three possible outcomes with three basis vectors
- Construct vector V by defining relationships with the basis set.

Third - We then construct a fourth unit vector, V. (See Figure 1) We define this vector by "orienting" it with respect to the basis set such that the square of the inner product of this fourth vector with each basis vector is equal to the probability for the toss outcome associated with that basis vector.

We therefore assign: $|(V , B1)|^2 = P1$, $|(V , B2)|^2 = P2$, and, $|(V , B3)|^2 = P3$.

The set of probabilities, by assigning the pairwise inner product relationships, fix the orientation of the fourth vector with respect to the three orthonormal basis vectors.

We can now proceed by highlighting particular points regarding the example and this methodology of using vectors to represent probability distributions.

4.3 Point 1: The L^2 – Probability Isomorphism

By far, the most important point to recognize is that this method of using vectors is different from the conventional way in which vectors are used in physical theories.

In previous physical theories, vectors are typically employed to represent physical phenomena that are characterized by a direction and a magnitude. In order to represent physical phenomena characterized in this way, one employs the norm of an individual vector and its angular relationship with some arbitrarily chosen reference vector.

Here, we see vectors being used in a different way. Here the vector structure is employed to represent a physical situation characterized by an observed probability distribution.

We note that it is a defining feature of probability distributions, that for any set of mutually exclusive exhaustive possible outcomes or transitions, the observed probabilities must sum to one. The L^2 vector norm of a unit vector displays this same structure.

Algebraic structure characterizing probabilities: 1 = p1 + p2 + …

Algebraic structure defining the L^2 vector norm (for a unit vector): $| V | = 1 = | (V , B1) |^2 + | (V , B2) |^2 + \ldots$ (where V is a unit vector, and $\{ B_n \}$ is an orthonormal basis set of n-dimensionality.)

Given this structural isomorphism, we can make pairwise assignments associating each probability in the sum with an inner product in the other sum, thereby representing one by using the other. We see that any possible probability distribution can be represented by assigning the inner product relationships of V with the basis vectors as necessary.

Note that this usage differs from the conventional by employing different structural features of the vectors and vector space. In this usage, it is the inner product relationship between a pair of unit vectors that actually does the representational work, ie., carries the information characterizing the physical situation. In particular, the vectors are used pairwise and not individually. A pair of vectors represents a pair of states (states of the die before and after the toss). The vector pairwise relationship represents a pairwise relationship that exists between the states. The relationship between the state pair is the scalar transition probability. The relationship between the vector pair is the scalar inner product modulus squared.

To note the substantial difference between the conventional use of vectors and this new use, note the different role of the norm of the vector. In conventional use, the "length" or norm of the vector plays a working role of representing the magnitude of the physical quantity. In quantum theory, however, every vector norm is assigned by the representational methodology to equal one. The norm carries no information characterizing the phenomena.

4.4 Point 2: Generalizability

The absolutely essential point prerequisite to understanding or explaining the quantum use of vectors is the recognition that the vectors are employed in the new and uniquely quantum way described above. From that, much of the particulars to be explained follow easily.

For example, seeing the fundamental representational principle at work in the example, it is apparent how it can be richly generalized. If the probability distribution to be represented, has n rather than three alternative possible transitions, then one selects a vector space of n dimension that has the required L^2 structure. The continuum case is also possible. In fact, it is only the L^2 structure doing representational work and so in any other respects, the vector space can be generalized. We have used real vectors to make the example intuitive. As long as the L^2 structure is retained, real, complex, or, more exotic vector spaces are usable in this way. The set of vector spaces generalized so as to retain the L^2 structure is the set of Hilbert spaces.

4.5 Point 3: Adopted Quantum Conventions

The example above employed a contrived situation using die tosses. The essential feature, however, is the presence of a probability distribution which one would like to represent mathematically. That we have the same requirement in quantum theory is apparent. For example, observations of spin component on a spin one object display the same set of three alternative possible transitions each with an observed probability.

We can see more explicitly, however, the theory's adoption of this convention by turning to the postulates. The theory adopts the following two:

P1: The state of an object is represented by a vector in a Hilbert space.

P2: The probability of a transition from one to state to some other state is given by the "Born Rule". The Born rule represents the transition probability as an inner product function on two vectors that represent the two physical states involved in the transition.

We see therefore, quite explicitly, that quantum theory adopts the representational method of using vectors demonstrated in the example above.

4.6 Point 4: Understanding R1 and R2

Recall from the beginning of this chapter, that our goal here is to understand the origins of the two requirements, R1 and R2, that played such an important role in the last chapter. That is, we want to have some intuition for the origins of the Born rule which is the heart of R1. As mentioned above, having explicitly and clearly acknowledged that the theory is using vectors in a new way, the remaining explanatory particulars follow easily. I provide again R1 and R2.

> R1: Vector, V, is subject to n independent constraints with respect to an orthonormal basis set of vectors $\{ b_i \}_{(i = 1 , n)}$. Each constraint is of the form, $p_i = | (V , b_i) |^2$.
> R2: Vector V is n-dimensional.

We now see the following regarding R1. The theory undertakes to represent a new class of physical phenomena not within the domain of classical physics, ie., probabilistic state transitions. That means the theory has an interest in mathematically representing probability distributions as a "best we can do" attempt to describe the state of an object. To do this, the quite reasonable choice is adopted to represent these probability distributions by exploiting the probability – L^2 structural isomorphism. Consequently, vectors are employed to represent states rather than phase space points. This commits the theory to representing a probability relating a state pair by using an inner product function on a vector pair. This is the origin of the Born rule that the theory adopts by postulate. In the explanation provided in the last chapter, what was done was to recognize that the Born rule is also, in fact, a pairwise vector constraint. It is this constraint situation that the theory relaxes by using complex vectors when required.

R2 is similarly understandable having recognized the new vector use method. From the simple example, we see clearly that the dimensionality of the vector space used is chosen specifically for the purpose of representing the set of all possible transitions and associated probabilities. The dimensionality, therefore, is assigned by the number of alternative transitions present.

In addition to the above comments on R1 and R2, having considered the origins of the quantum representational postulates, we now see clearly why in the last chapter, our beginning point was essentially the option of using either real or complex vectors. This is a consequence of the theory's fundamental commitment to the use of a vector space with the L^2 structure. The L^2 requirement is a strong requirement fixing much of the structure of the vector space. In the present case, the L^2 requirement tells us that if we require two real variables per vector dimension, then those two must obey the complex number combination rule.

4.7 Point 5: Abstracting the Methodology

Just as a point of interest, there is a useful and more general way to characterize the new vector use methodology.

What is being represented by this method is the pairwise relationships that are empirically observed between the possible states of an object. An observed probability for a transition from one state to another defines a scalar relationship between that pair. Since every pairing of possible states has some observed transition probability, then that defines a metric structure of scalar pairwise relationships on the entire space of possible states.

Adoption of the Born rule adopts a methodology of representing pairwise state scalar relationships by using pairwise vector scalar relationships. The vector scalar relationship is an inner product function on the vector pair associated with the states.

What we recognize is this. The physics drives the math. The metric structure of the state space is a brute physical fact. If we choose to represent this metric structure by using the metric structure defined by inner products on vector pairs, then this choice imposes a variety of hard requirements on the metric structure of the vector space.

The complex vector requirement in the representational space is a consequence of one particular requirement which originates in the state space. It is an issue of pairwise relationship freedom/constraint. The state space displays multiple interrelated pairwise scalar relationships. The representational vector space must provide sufficient structural freedom to represent these same pairwise scalar relationships. Normally, to provide additional required freedom, if needed, one would be tempted to increase the dimensionality of the vector representational space. (For example try orienting vectors a, b, and c, in a two dimensional real vector space at angle(a,b) = 15, angle(a,c) = 30, and angle(a,c) = 20. We can do it, but, by going to three dimensions.) This is not an option, however, since the methodology assigns a dimensionality equal to the number of alternative possible transitions. Since the representational vector space must meet the hard representational demands originating in the state space, we turn to a second alternative avenue for including additional variables to relax constraints. That is, rather than increasing dimensionality, one increases the number of variables per dimension.

4.8 Chapter Concluding Comments

A main point of this chapter is to provide some insight as to the origins of requirements R1 and R2, which played important roles in the explanation provided in the previous chapter.

We could have left them as simply matters of convention adopted by the theorist during theory construction. They are not, however, impenetrable quantum doctrine and are, in fact, adopted for good reason. This chapter extends the explanation from the last chapter beyond the postulates and, hopefully, makes good connection with the underlying physical and representational principles.

CHAPTER 5

SUMMARY COMMENTS

The main purpose of the dissertation was to provide an explanation for the role of complex numbers in quantum theory.

I have explained that the point of entry of the complex requirement into the theory is in the construction of the vectors which represent states of objects. In order to explain why these vectors are complex, it was necessary to recognize that the theory makes a commitment to the use of vectors to represent states that is more fundamental than the theory's commitment to the use of complex vectors. That is, to ask why the theory requires complex numbers is, in a sense, the wrong question. In this case, the right question is, given that this methodology of using vectors is to be employed to represent states, then why must those vectors be complex vectors?

Turning to the state vectors, we recognize that quantum theory adopts a method of using vectors to represent physical phenomena that is different from their use in previous theories. It is attention to this new quantum representational methodology, adopted by postulate, that then leads quite directly to an explanation of the complex vector requirement.

I have shown, in particular, that the fundamental issue that brings the complex vector requirement is an issue of freedom versus constraint. The theory represents pairwise relationships between states of objects by using pairwise relationships between vectors. It was a key point to recognize that this representation also constitutes a constraint on the relationship between the two vectors. Consequently, this method of using vectors can only go so far. Assigning multiple interrelated pairwise vector relationships may not be possible. The structure of the vector space may not allow it. So we see that the vector use method faces the risk of vector space structural limitations. It is in this issue, freedom versus constraint where the choice of vectors used, real or complex, makes a difference. Complex vectors, quite simply, provide more available independent variables than do real vectors. More variables translate into more freedom in the vector pairwise assignments. It is for this reason, to relax a pairwise vector relationship constraint, that quantum theory requires complex vectors.

If quantum theory requires complex numbers to relax a constraint situation, then it does not require them for the other, sometimes suggested, reasons. The complex numbers are not present in the theory to represent or reflect some complex feature of the observed physical phenomena. That is, nature is not complex. Similarly, the complex numbers are not present as a mathematical convenience, ie., to facilitate some mathematical manipulation.

In the dissertation I have made a point to show the complex number issue as a secondary issue located in the more general context of the theory's adopted convention for the use of vectors to represent states. The real issue here and the real insight into the theory is the recognition of the new quantum principle of vector use. Quantum theory has a different theoretical task than classical. Quantum theory undertakes to represent physical phenomena to include probabilistic state transitions. The theory needs, therefore, to mathematically represent probability distributions. The new use of vectors is a diabolically clever method of doing just that. To exploit the L^2 structure of the vector space rather than other features would be a stroke of genius if done intentionally by a theorist rather than, what seems more likely, by fortunate historical happenstance.

I hope that you, the reader, have found the discussion of interest. Thank you.

REFERENCES

[1] Chen Ning Yang, Square Root of Minus One, Complex Phases, and Erwin Schrodinger, published in Kilmister, C. W., Ed., *Schrodinger, Centenary Celebration of a Polymath*, Cambridge University Press, (2010), p56

[2] Wheeler, J. A., Trans. N. Y. Academy of Science, 33, (1971), 745; cf. p. 778

[3] Aharonov, Y., *Quantum Paradoxes*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, (2005)

[4] Goyal, P., Knuth, K. H., Skilling, J., Origin of Complex Quantum Amplitudes and Feynman's Rules, Phys. Rev. A 81, (2010)

[5], Schleich, W. P., Greenberger, D. M., Kobe, D. H., and Scully, M. O., Schrödinger equation revisited**,** PNAS 2013 110 (14), (2013), 5374-5379

[6] Stückelberg E, Quantum theory in real Hilbert space, Helv Phys Acta **33**, (1960), 727–752.

[7] Wheeler JA It from bit, (1990) published in Kobayashi, S., Ed., *Foundations of Quantum Mechanics in the Light of New Technology*, Physical Society of Japan, Tokyo, pp 354–368.

[8] Schleich WP, *Quantum Optics in Phase Space*, Wiley–VCH, Weinheim, Germany, (2001)

[9] Bohm, D., Phys. Rev. 85, (1952), 166-179

[10] Bohm, D., Phys. Rev. 85, (1952), 180-193

[11] Holland, P. R., *The Quantum Theory of Motion*, Cambridge University Press, (2000)

[12] Maynard, G., Lambert, D., Deering, W. D., The Requirement for Complex Numbers in Quantum Theory, Physics Education Journal (IAPT), Volume 31, Issue 2 (2015)