The Requirement for Complex Numbers in Quantum Theory

G. Maynard[1] and D. Lambert[1] and W. D.Deering[1]

[1]Department of Physics
University of North Texas
Denton, TX USA.

(*Submitted*13-04-2015)

Abstract

Generations of physicists have at some point wondered about the role of complex numbers in quantum theory. None have had this point explained. Taking this as a pedagogical issue, this article elucidates the origins of the complex requirement and draws the connection between these origins and the requirement that the state vector be complex.

1. Introduction

A recent article by Sivakumar[1] highlights an important pedagogical void which has long plagued the teaching of introductory and fundamental principles of quantum theory. That is, there is no generally accepted explanation available for the apparent requirement for complex numbers in the mathematical formulation of the theory. Sivakumar provides a simplified demonstration (due to Sakurai[2] and Townsend [3]) that, in fact, the complex numbers are required. In this article, we continue beyond the demonstration and attempt to explain the underlying issues.

Quantum theory employs unit vectors to mathematically represent states of physical objects. In the following, we will identify four requirements which must be satisfied by these vectors. We then show that the four requirements are not satisfied by real vectors, but can be satisfied by complex vectors.

In the next section, we begin by simply stating the four requirements. We then devote an individual section to each requirement and discuss the physical and theoretical origins of that requirement.Turning from origins, we then consider them simply as a set of requirements on vector structure. By identifying this set of specific requirements on the vectors, we see clearly at exactly what point and for exactly what reason real vectors fail to satisfy the requirements.

We note up front that the article intent is pedagogical. Accordingly, points are presented in what is, hopefully, an intuitive, and conceptual way. We ask some leeway in completeness and rigor.

We also note that some closely related issues are put aside. For example, why the theory adopts use of vectors and the Born rule as a representational convention is an important foundational question. Here, we accept as a starting point that the theory does this. By making these choices, we focus on the question of, assuming that vectors are to be used in this way, why must the vectors be complex.

2. The Requirements

As mentioned, the theory uses vectors to represent states of physical objects. In particular, we will be interested in the representation of angular momentum states. It is in representing these states that the complex requirement arises.

The following four requirements must be satisfied by any vector, V, used by the theory to represent an angular momentum state. We will show that they are not satisfied by real vectors.

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, $p_i = p_i(r)$

In the next sections, we discuss these requirements individually with emphasis on their origins.

3. Requirement R1

In this section we state and discuss requirement R1.

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, $p_i = |(V , b_i)|^2$.

Discussion:

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 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 probability value. From this perspective, the Born Rule, in fact, defines the vector pair (a partial definition) by specifying their relation. 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.

We 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 the L^2 structure.

4. Requirement R2

In this section we state and discuss requirement R2.

<u>R2: The vector is n dimensional</u>.

Discussion:

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.

5. Requirement R3

In this section we state and discuss requirement R3.

<u>R3: Vector V must vary with two real variables, V (r, c</u>).

Discussion:

It is an 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[1] will yield different results. 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, 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" 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 will be useful to choose a different coordinatization. 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 .

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. We 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.

6. Requirement R4

In this section we state and discuss requirement R4.

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.

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, $p_i = |(V , b_i)|^2$)

Discussion:

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

[1] The difference involves probabilities and consequently is observed on ensembles of similarly prepared objects.

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".

7. Satisfying The Requirements

In this section we collect again the four requirements and show that they are not satisfied by real vectors, but can be by complex vectors.

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.

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, $p_i = p_i(r)$

We show first that real vectors do not satisfy these requirements as follows:

Point 1: Number of variables present

Assume V is real. 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:

1. Is fully specified by "r", and
2. 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.

We 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.

8. Discussion

Hopefully the analysis presented in this article is of pedagogical value. We have separated out issues in order to provide good access to the role of complex number in the theory.

We have identified a set of four specific requirements on the vectors employed by the theory. The goal was to facilitate two different perspectives linked by this set of requirements.

One perspective is mathematical. One can disregard the origins of the requirements and consider them simply as given. From this starting point, the exercise is one of vector structure. One can observe the interplay of freedom and constraint considerations that prevent the requirements from being satisfied by real vectors.

The other perspective is physical.Here we disregard the mathematical implications, and trace back the origins of the requirements. What specific features of the physical phenomena or adopted theoretical conventions impose these requirements?

The set of four requirements therefore serves as a point on which to stand and contemplate both available perspectives. From there, the student of foundational quantum theory can find a traceable connection all the way from the physical and theoretical origins through to their end consequence, a particular mathematical detail in the formal theory, the presence of complex numbers. More importantly, the student has a useful framework to separate out issues and make their own evaluation of the requirements, their origins, and their implications.

We point out two particular results of our analysis.

It is sometimes commented in the foundational literature that to explain some particular mathematical detail of the theory would be to elucidate its physical origins. Here we see that there is an identifiable physical origin. The complex requirement is, in part, a consequence of the fact that angular momentum states vary in two physical space orientation degrees of freedom. There is, however, a second equally important origin. It is the theory's adoption of a particular representational convention that imposes very substantial pairwise constraints on the vectors employed. Thus we see that, in this case, elucidating physical origins is not sufficient. We also must elucidate the theoretical representational conventions adopted.

We also mention a second result. We now have an answer to the big question, why are complex numbers required by the theory? The fundamental reason that they are required is to resolve a disappointingly mundane issue of freedom versus constraint. Quite simply, they provide more variables than real vectors. State vectors are subject to the significant "Born Constraints" yet must also honor a freedom demand when representing angular momentum states. The vectors must satisfy both. As we have seen, real vectors come up short on available variables. Consequently, we find complex vectors employed to represent states of physical objects.

References :

[1] S.Sivakumar, Physics Education, **28**No3.(2012)
[2] J. J. Sakurai, Introduction to Modern Quantum Mechanics (Addison-Wesley 1994, New York) p27.
[3] J. S. Townsend, A Modern Introduction to Quantum Mechanics (McGraw Hill, 1992, Singapore) p17.