

# The Origin of Complex Quantum Amplitudes

Philip Goyal\*, Kevin H. Knuth<sup>†</sup> and John Skilling\*\*

\*Perimeter Institute, Waterloo, Canada. Email: pgoyal@perimeterinstitute.ca

<sup>†</sup>University of Albany, NY, USA. Email: kknuth@albany.edu

\*\*Maximum Entropy Data Consultants Ltd, Kenmare, Ireland. Email: skilling@eircom.net

**Abstract.** Physics is real. Measurement produces real numbers. Yet quantum mechanics uses complex arithmetic, in which  $\sqrt{-1}$  is necessary but mysteriously relates to nothing else. By applying the same sort of symmetry arguments that Cox [1, 2] used to justify probability calculus, we are now able to explain this puzzle.

The dual device/object nature of observation requires us to describe the world in terms of *pairs* of real numbers about which we never have full knowledge. These pairs combine according to complex arithmetic, using Feynman's rules.

**Keywords:** Feynman rules, complementarity.

**PACS:** 03.65.Ta

## INTRODUCTION

Measurement always involves an *interaction* between observed object and observing device. Presumably, the device returns to us only *one* of a *pair* of numbers that quantified the interaction. We do not probe the inaccessible detail of the interaction — plausibly there may be no classical model for it. It suffices to note that we never attain complete knowledge of either the object (which interacted with the imperfectly known device) or the device (which interacted with the imperfectly known object). We can never bootstrap our way to total knowledge [3], and this indicates that our knowledge is doomed to be, at least in part, probabilistic.

To set the scene, recall that the laws of probability calculus are firmly founded on elementary symmetries. Propositions can be combined under logical OR, with its subsidiary dual AND, to form a distributive lattice. Associativity of OR,

$$(X \text{ OR } Y) \text{ OR } Z = X \text{ OR } (Y \text{ OR } Z)$$

requires any valuation  $\mu$  to obey the sum rule  $\mu(X \text{ OR } Y) + \mu(X \text{ AND } Y) = \mu(X) + \mu(Y)$ . Distributivity of AND over OR, and of OR over AND,

$$(X \text{ OR } Y) \text{ AND } Z = (X \text{ AND } Z) \text{ OR } (Y \text{ AND } Z)$$

$$(X \text{ AND } Y) \text{ OR } Z = (X \text{ OR } Z) \text{ AND } (Y \text{ OR } Z)$$

ensures that the values on the elementary propositions (that cannot be further decomposed) can be set arbitrarily. Ordering (that  $X$  is more precise than  $(X \text{ OR } Y)$ ) ensures that these values  $\mu(X)$  are non-negative — this is the basis of measure theory [4, 5].

Probability  $\Pr(X | I)$  is a bi-valuation referring to proposition  $X$  in context  $I$ . Ordering of context (that  $X$  is identified less ambiguously within  $I$  than in the wider

( $I$  OR  $J$ ) requires the bi-valuation to have the form of a scaled measure, in the ratio form  $\Pr(X | I) = \mu(X \text{ AND } I) / \mu(I)$  which encapsulates the ordinary laws (sum rule, product rule and  $(0, 1)$  range) of probability. There is no choice about the calculus of inference. We claim that the basic quantum calculus of physics is similarly forced by general desiderata, without input from classical physics.

## OPERATIONAL FRAMEWORK

Knowledge of an object comes from measurements  $\mathbf{M}$ , which have outcomes  $m$ . We denote a sequence of measurements or outcomes by square brackets [6]. Given an initial measurement  $\mathbf{M}_1$  with outcome  $m_1$ , the probability of a sequence of outcomes  $A = [m_1, m_2, \dots, m_k]$  of measurements  $[\mathbf{M}_1, \mathbf{M}_2, \dots, \mathbf{M}_k]$  is defined as

$$\Pr(A) = \Pr(m_k, \dots, m_2 | m_1), \text{ given } \mathbf{M}_k, \dots, \mathbf{M}_2, \mathbf{M}_1. \quad (\text{Probability})$$

For simplicity, we choose to consider *adjacent* measurements, taken at times sufficiently close so as to exclude significant external interaction, as well as significant spontaneous evolution of the object. An object may have several features, but we choose to consider measurements that exhibit closure [7]. *Closure* means that successive measurements of some feature form a Markov chain, with future outcomes influenced by the present, but not by the past. Thus, successive measurements  $[\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3, \dots]$  (possibly different) have outcomes  $[m_1, m_2, m_3, \dots]$  for which the sequence probability is

$$\Pr(m_{k+1}, m_k, m_{k-1}, \dots | m_1) = \Pr(m_{k+1} | m_k), \text{ given } \mathbf{M}_{k+1}, \mathbf{M}_k, \dots, \mathbf{M}_1. \quad (\text{Closure})$$

We assert that if the *same* measurement  $\mathbf{M}$  is repeated adjacently, then successive outcomes  $m, m'$  are the same, so that

$$\Pr(m' | m) = \delta_{mm'} \text{ given } \mathbf{M}. \quad (\text{Repeatability})$$

*Repeatability* ensures that objects can have observational permanence, and can be seeded by selecting an initial outcome.

A *coarse-grained* measurement groups the outcomes into equivalence classes. For example, while measurement  $\mathbf{M}$  may have three outcomes  $m, m'$  or  $m''$ , a coarse-grained measurement  $\tilde{\mathbf{M}}$  might only have two outcomes  $m$  and  $(m', m'')$  where  $m'$  and  $m''$  are not distinguished, or even only one outcome  $(m, m', m'')$ . We say that the coarse measurement  $\tilde{\mathbf{M}}$  can be *refined* by performing the measurement  $\mathbf{M}$ . A measurement is called *atomic* when it cannot be further refined.

## SYMMETRIES

Probability theory quantifies propositions, related in a lattice whose symmetries define the calculus. Likewise, in the operational view of physics, we have *sequences of outcomes* from a program of measurements, related by *parallel* and *series* combinations, whose symmetries will — subject to a subtle connection with probability — define the calculus.

## Sequences in parallel

Suppose that program  $[\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3]$  is set up. On one run, the sequence of outcomes is  $A = [m_1, m_2, m_3]$ , but on another run the sequence is  $B = [m_1, m'_2, m_3]$ , differing ( $m_2 \neq m'_2$ ) in just one outcome. Now consider a second program, identical to the first except that the differing outcome is measured coarsely by  $\tilde{\mathbf{M}}_2$ , which does not distinguish  $m_2$  from  $m'_2$ . This program can generate an outcome  $C = [m_1, (m_2, m'_2), m_3]$ . We say that  $C$  combines  $A$  and  $B$  *in parallel* (Fig. 1), and we write

$$C = A \vee B. \quad (\text{Parallel})$$

Note that there could be several different meanings of  $\vee$  depending on which measurement is coarsened, but we will be consistent and only use one at a time.

Because  $(m_2, m'_2)$  is the same as  $(m'_2, m_2)$ , it follows that  $[m_1, (m_2, m'_2), m_3] = [m_1, (m'_2, m_2), m_3]$ , so that  $\vee$  is *commutative*.

$$A \vee B = B \vee A \quad (1)$$

Now consider three sequences  $A = [m_1, m_2, m_3]$ ,  $B = [m_1, m'_2, m_3]$  and  $C = [m_1, m''_2, m_3]$ , all different because outcomes  $m_2, m'_2, m''_2$  all differ. These sequences can be combined into  $[m_1, (m_2, m'_2, m''_2), m_3]$  in two different ways, namely  $(A \vee B) \vee C$  and  $A \vee (B \vee C)$ , implying that  $\vee$  is also *associative*.

$$(A \vee B) \vee C = A \vee (B \vee C) \quad (2)$$

## Sequences in series

Suppose outcomes  $C = [m_1, m_2, m_3]$  were obtained from program  $[\mathbf{M}_1, \mathbf{M}_2, \mathbf{M}_3]$ . Because of repeatability, repeated application of the measurement  $\mathbf{M}_2$  would keep giving the same outcome  $m_2$ . Thus the sequence  $C = [m_1, m_2, m_3]$  is equivalent to  $A = [m_1, m_2]$  followed by  $B = [m_2, m_3]$ . Concatenation, with the last measurement and outcome of the earlier sequence coinciding with the first measurement and outcome of the later, is called combination *in series* (Fig. 2), and we write this non-commutative operation as

$$C = A \bullet B. \quad (\text{Series})$$

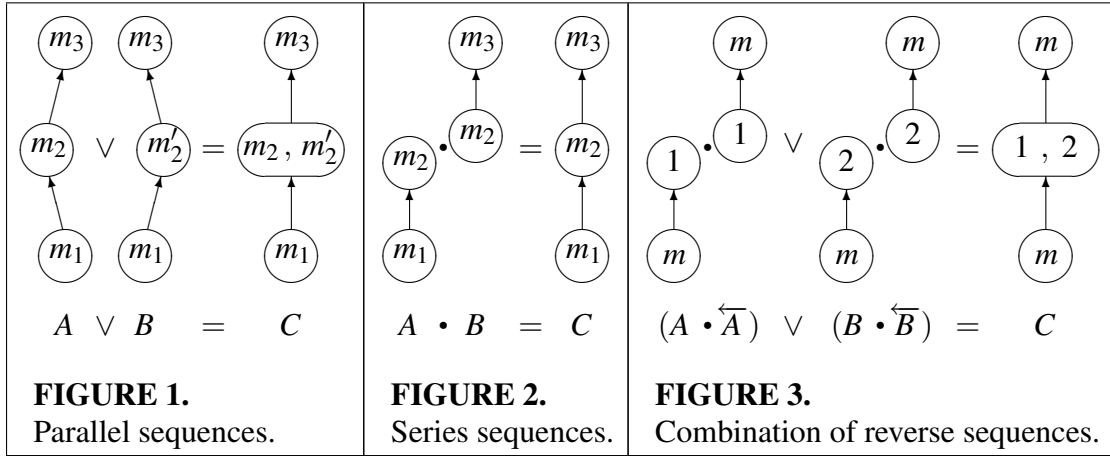
By chaining three sequences  $A = [m_1, m_2]$ ,  $B = [m_2, m_3]$  and  $C = [m_3, m_4]$ , we see that  $\bullet$  is *associative*.

$$(A \bullet B) \bullet C = A \bullet (B \bullet C) \quad (3)$$

Now consider sequences  $A = [m_1, m_2, m_3]$  and  $B = [m_1, m'_2, m_3]$ , each concatenated with  $C = [m_3, m_4]$ . These can be combined into  $[m_1, (m_2, m'_2), m_3, m_4]$  in two different ways, namely  $(A \vee B) \bullet C$  and  $(A \bullet C) \vee (B \bullet C)$ , implying that  $\bullet$  is *right-distributive* over  $\vee$ . Similarly, it is also *left-distributive*.

$$(A \vee B) \bullet C = (A \bullet C) \vee (B \bullet C) \quad \text{and} \quad C \bullet (A \vee B) = (C \bullet A) \vee (C \bullet B) \quad (4)$$

This completes the list (1,2,3,4), identified in [8], of symmetries obeyed by sequences.



## SEQUENCE PAIRS

In probability theory, identification of the abstract symmetries of a lattice is followed by the specification of a representation for its elements. Within fixed context, those elements are propositions having partial order, whereupon transitivity of ordering shows that no loss of generality is involved in letting the representation use inherently-transitive real scalar numbers.

For sequences with fixed first and last outcomes, there is no similar partial order requiring scalar representation. Instead, we take our cue from the object/device dual nature of a measurement, and adopt the *pair postulate*.

**Pair Postulate:** each sequence of outcomes from a given measurement program is represented by a *pair* of real numbers, with the probability of this sequence being a continuous, non-trivial function of both components of this real number pair.

Sequence  $A$  is represented by a pair  $\mathbf{a} = (a_1, a_2)^\top$ , which is to determine its probability through some function  $p$ ,

$$\Pr(A) = p(\mathbf{a}).$$

Our task is to identify a calculus for pairs, in which parallel and series combinations of outcomes can be computed, and to complete the job by determining the function  $p$ . Parallel and series operations on sequences are to be represented by corresponding operators, pair-valued and continuous, on pairs.

$$A \vee B \text{ is represented by } \mathbf{a} \oplus \mathbf{b}, \quad \text{and} \quad A \cdot B \text{ is represented by } \mathbf{a} \odot \mathbf{b}.$$

The symmetries obeyed by sequences, namely commutativity and associativity of  $\vee$ , and associativity and distributivity of  $\cdot$ , must be mirrored by relationships obeyed by

the operators, so that

$$\begin{aligned}
\mathbf{a} \oplus \mathbf{b} &= \mathbf{b} \oplus \mathbf{a} \\
(\mathbf{a} \oplus \mathbf{b}) \oplus \mathbf{c} &= \mathbf{a} \oplus (\mathbf{b} \oplus \mathbf{c}) \\
(\mathbf{a} \odot \mathbf{b}) \odot \mathbf{c} &= \mathbf{a} \odot (\mathbf{b} \odot \mathbf{c}) \\
(\mathbf{a} \oplus \mathbf{b}) \odot \mathbf{c} &= (\mathbf{a} \odot \mathbf{c}) \oplus (\mathbf{b} \odot \mathbf{c}) \\
\mathbf{c} \odot (\mathbf{a} \oplus \mathbf{b}) &= (\mathbf{c} \odot \mathbf{a}) \oplus (\mathbf{c} \odot \mathbf{b})
\end{aligned}$$

## Sum and product rules

The vector associativity theorem [9] states that the continuous, commutative and associative operator  $\oplus$  admits a continuous, invertible function  $\mathbf{F}$  such that

$$\mathbf{F}(\mathbf{a} \oplus \mathbf{b}) = \mathbf{F}(\mathbf{a}) + \mathbf{F}(\mathbf{b}).$$

Hence we can without loss of generality work in the transformed pair-space  $\mathbf{F}(\cdot)$  where  $\oplus$  is component-wise addition (the *sum rule*).

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \oplus \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \end{pmatrix}$$

Effectively the same derivation underlies the scalar sum rule of probability calculus, where the sum rule is insufficient to fix the scale. Here, the remaining freedom is the linear transformation

$$\begin{pmatrix} c'_1 \\ c'_2 \end{pmatrix} = \begin{pmatrix} T_1 & T_2 \\ T_3 & T_4 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

with  $T_1T_4 - T_2T_3 \neq 0$  to retain invertibility.

We proceed to show that  $\odot$  is a form of multiplication. Left and right distributivity of  $\odot$  shows the operator to be linear in both arguments, so that it must take the multiplicative form

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \odot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = a_1b_1 \begin{pmatrix} \gamma_1 \\ \gamma_5 \end{pmatrix} + a_1b_2 \begin{pmatrix} \gamma_2 \\ \gamma_6 \end{pmatrix} + a_2b_1 \begin{pmatrix} \gamma_3 \\ \gamma_7 \end{pmatrix} + a_2b_2 \begin{pmatrix} \gamma_4 \\ \gamma_8 \end{pmatrix}$$

where  $\gamma = (\gamma_1, \gamma_2, \gamma_3, \gamma_4; \gamma_5, \gamma_6, \gamma_7, \gamma_8)$  involves 8 universal constants, initially arbitrary. Associativity of  $\odot$  restricts  $\gamma$  to forms that can be further simplified by using the freedom to transform linearly. Three standard forms result — the analysis is lengthy and given elsewhere [10] — along with some degenerate alternatives that fail to capture the full freedom of pairs and give nothing useful:

$$\gamma = (1, 0, 0, \mu; 0, 1, 1, 0) \quad \text{with } \mu = -1 \text{ or } 0 \text{ or } 1.$$

Each of these is an acceptable definition of multiplication, as it happens commutative.

## Sequence probability

We now analyze the relation between pairs and probability, because according to the pair postulate, the calculus of pairs should admit a function  $p(\mathbf{a})$  that can be identified

with probability. Consider successive sequences  $A = [m_1, m_2]$  and  $B = [m_2, m_3]$ , that combine to  $C = A \bullet B = [m_1, m_2, m_3]$ . The probability associated with  $C$  is in general

$$\Pr(m_3, m_2 \mid m_1) = \Pr(m_3 \mid m_2, m_1) \Pr(m_2 \mid m_1)$$

but when we restrict to measurements with closure the history vanishes and we get

$$\Pr(C) = \Pr(m_3, m_2 \mid m_1) = \Pr(m_3 \mid m_2) \Pr(m_2 \mid m_1) = \Pr(B) \Pr(A)$$

Corresponding to this, the pair-probability function  $p$  must obey

$$p(\mathbf{a} \odot \mathbf{b}) = p(\mathbf{a}) p(\mathbf{b})$$

This equation can be solved [10] for all three forms of multiplication, giving

$$p(\mathbf{a}) = \begin{cases} (a_1^2 + a_2^2)^{\alpha/2} & \text{for } \mu = -1, \\ |a_1|^\alpha e^{\beta a_2/a_1} & \text{for } \mu = 0, \\ |a_1|^\alpha |a_2|^\beta & \text{for } \mu = +1, \end{cases}$$

where  $\alpha$  and  $\beta$  are universal constants.

## Reverse sequences

If program  $[\mathbf{M}, \mathbf{N}]$  produces sequence  $A = [m, n]$ , then the reverse program  $[\mathbf{N}, \mathbf{M}]$  producing sequence  $\overleftarrow{A} = [n, m]$  ought to be related. There are, after all, situations where geometrical transformation could turn  $\mathbf{M}$ -giving- $m$  into  $\mathbf{N}$ -giving- $n$ , and vice versa. Accordingly, we assume an invertible relationship

$$\overleftarrow{\mathbf{a}} = \mathbf{R}(\mathbf{a})$$

between the pair  $\mathbf{a}$  associated with  $A$  and the pair  $\overleftarrow{\mathbf{a}}$  associated with the reverse  $\overleftarrow{A}$ . Parallel combination  $\overleftarrow{C} = \overleftarrow{A} \vee \overleftarrow{B}$  reverses to  $\overleftarrow{C} = \overleftarrow{A} \vee \overleftarrow{B}$ , while series combination  $C = A \bullet B$  reverses to  $\overleftarrow{C} = \overleftarrow{B} \bullet \overleftarrow{A}$ . Hence

$$\mathbf{R}(\mathbf{a} + \mathbf{b}) = \mathbf{R}(\mathbf{a}) + \mathbf{R}(\mathbf{b}) \quad \text{and} \quad \mathbf{R}(\mathbf{a} \odot \mathbf{b}) = \mathbf{R}(\mathbf{b}) \odot \mathbf{R}(\mathbf{a}).$$

From the additivity, the reversal operator  $\mathbf{R}$  has to be a linear transformation, whose precise definition depends on the chosen form of multiplication  $\odot$ .

We use  $\mathbf{R}$  to analyze the situation where program  $[\mathbf{M}, \tilde{\mathbf{N}}, \mathbf{M}]$  produces sequence  $C = [m, (1, 2), m]$ , in which 1 and 2 are the only possible outcomes of the atomic measurement  $\mathbf{N}$  coarsened to  $\tilde{\mathbf{N}}$ . Writing  $A = [m, 1]$  and  $B = [m, 2]$  as the possible sequences for program  $[\mathbf{M}, \mathbf{N}]$  seeded as  $m$ , we have (Fig. 3)

$$C = (A \bullet \overleftarrow{A}) \vee (B \bullet \overleftarrow{B}).$$

As for probabilities,  $A$  and  $B$  are exclusive and exhaustive, so  $p(\mathbf{a}) + p(\mathbf{b}) = 1$ . Moreover,  $\tilde{\mathbf{N}} = (1, 2)$  is the trivial measurement, so the final outcome of  $C$  is forced to equal the initial outcome, meaning that  $C$  is certain, so

$$p(\mathbf{a}) + p(\mathbf{b}) = 1 \quad \implies \quad p(\mathbf{c}) = 1.$$

This requirement eliminates the second and third forms. Only the first survives, with  $\alpha = 2$  and  $\mathbf{R}(\mathbf{a}) = (a_1, -a_2)^\top$ . Proof of this is in [10]. Our task is done.

## SUMMARY

We describe physics operationally through sequences of measurements that are quantified by *pairs* of real numbers. This is motivated by the dual (device/object) nature of measurement, and formalized through the *pair postulate* which notes that our knowledge of a pair is always incomplete, hence probabilistic. Our postulate is a formal expression of quantum “complementarity” [11, 12].

Sequences of measurement obey symmetries. Commutativity and associativity of “parallel” require the sum rule.

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \oplus \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \end{pmatrix}$$

Associativity and distributivity of “series” allows a choice of three product rules.

A pair  $\mathbf{c}$  becomes observable through the probability  $p(\mathbf{c})$  associated with it. Applying this requirement in the Markovian case of measurements with closure imposes a particular form of  $p$  for each product rule. Consideration of reverse sequences completes the specification, and gives

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \odot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 b_1 - a_2 b_2 \\ a_1 b_2 + a_2 b_1 \end{pmatrix} \quad \text{with} \quad p(\mathbf{c}) = c_1^2 + c_2^2.$$

These are the *Feynman rules*, applicable generally. Pairs are known as *quantum amplitudes* and behave as complex numbers, adding in parallel and multiplying in series, with modulus-squared giving the observable probability. We now see why quantum mechanics uses complex numbers. Quantification is “really” in terms of real pairs, but these behave like single complex entities.

## ACKNOWLEDGMENTS

Philip Goyal would like to thank Janos Aczél, Bob Coecke, Yitong Fu, and Luca Mana for very helpful discussions, and Perimeter Institute for excellent institutional support. Research at Perimeter Institute is supported in part by the Government of Canada through NSERC and by the Province of Ontario through MEDT. Kevin Knuth would like to thank Carlos Rodríguez and Ariel Caticha for many insightful discussions. John Skilling thanks the Perimeter Institute for hospitality.

## REFERENCES

1. R. T. Cox, *Am. J. Phys.* **14**, 1–13 (1946).
2. R. T. Cox, *The Algebra of Probable Inference*, The Johns Hopkins Press, 1961.
3. P. Rudolfsson, Private communication (2009).
4. K. H. Knuth, “Deriving Laws from Ordering Relations,” in *Bayesian Inference and Maximum Entropy Methods*, edited by G. J. Erickson, and Y. Zhai, AIP, New York, 2003, vol. 707, pp. 204–235.
5. K. H. Knuth, “Measuring on Lattices,” in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Oxford MS, USA, 2009*, edited by P. Goggans, AIP, New York, 2009.
6. A. Caticha, *Phys. Rev. A* **57**, 1572 (1998).
7. P. Goyal, *Phys. Rev. A* **78**, 052120 (2008).
8. Y. Tikochinsky, *Int. J. Theor. Phys.* **27**, 543 (1988).
9. J. Aczél, and M. Hosszú, *Acta Mathematica Hungarica* **7**, 327 (1956).
10. P. Goyal, K. Knuth, and J. Skilling (2009), arXiv:0907.0909.
11. N. Bohr, *Philosophy of Science* **4**, 289–298 (1937).
12. W. Heisenberg, *Zs. Phys.* **33**, 879–893 (1925), translation in [13].
13. B. L. van der Waerden, editor, *Sources of Quantum Mechanics*, Dover Publications, 1967.