Preprints.org

Article

Not peer-reviewed version

An Axiomatic Theory of Quaternionic Probability Extending Kolmogorov's Framework

Carlos Riveros Berger *

Posted Date: 28 September 2025

doi: 10.20944/preprints202509.2314.v1

Keywords: quaternionic probability; noncommutative probability; Kolmogorov axioms; quaternionic random variables; conditional probability; quaternionic Markov chains; quaternionic information theory; Clifford algebras; quaternionic entropy; quaternionic Bayesian inference

Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

An Axiomatic Theory of Quaternionic Probability Extending Kolmogorov's Framework

Carlos Riveros Berger

Program in Materials Science and Engineering, University of Santiago, Chile (USACH), Santiago, Chile; carlos.riverosb@usach.cl

Abstract

We present an axiomatic construction of quaternionic probability, extending Kolmogorov's classical framework to the noncommutative algebra of quaternions. The theory introduces quaternionic probability spaces, conditional probabilities, Bayes' rules, independence, random variables, expectations, and transport equations, all formulated in a consistent manner. Classical probability is recovered through scalar projection, while restriction to complex subalgebras reproduces the standard quantum formalism. Uniquely quaternionic structures arise, including noncommutative conditional probabilities, inequivalent forms of independence, and quaternionic transport laws. The framework further develops quaternionic Markov chains, entropy, and divergence measures that separate scalar uncertainty from vectorial coherence. Several illustrative examples are provided to show how quaternionic probability captures order effects, hidden correlations, and orthogonal divergences—features invisible to both classical and complex approaches. These results establish quaternionic probability as a rigorous generalization of Kolmogorov's axioms and as a potential foundation for future studies in noncommutative probability, integrable structures, and quaternionic extensions of mathematical physics.

Keywords: quaternionic probability; noncommutative probability; Kolmogorov axioms; quaternionic random variables; conditional probability; quaternionic Markov chains; quaternionic information theory; Clifford algebras; quaternionic entropy; quaternionic Bayesian inference

1. Introduction

The foundations of modern probability theory were established by Kolmogorov [1], and subsequently developed in great detail in classical texts such as Billingsley [2], Doob [3], Feller [4], and Parthasarathy [5]. This framework, based on real-valued measures on sigma-algebras, has proven remarkably successful in describing a wide variety of stochastic phenomena. Nevertheless, over the past decades there has been increasing interest in extending the notion of probability beyond the real line, in order to model systems where interference, order-dependence, or non-commutativity play a fundamental role.

Already in the mid-20th century, Feynman speculated on the role of *negative probabilities* in quantum theory [6], while Khrennikov and others explored generalized frameworks for probability, including non-Kolmogorovian and ppp-adic formulations [7,8]. In parallel, noncommutative probability theory emerged in operator algebras, with Accardi [9], Accardi–Cecchini [10], Voiculescu [11], Speicher [12], Wysoczański [13], and Kümmerer [14] among the leading contributors. These works highlighted that probabilistic notions such as independence and conditional expectation admit multiple, non-equivalent generalizations once the commutativity of real-valued measures is abandoned.

Quaternions, introduced by Hamilton in the 19th century [15], have also played a fundamental role in mathematical physics. Quaternionic analysis [16], quaternionic quantum mechanics [17–19], and quaternionic differential operators [20] provided natural frameworks for describing rotations, spin, and field theories where complex numbers are not sufficient. More recently, Danielewski and

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

Jadczyk [21] have revisited the foundations of quaternionic quantum mechanics, and Colombo, Sabadini, and Struppa [22] developed a functional calculus for slice hyperholomorphic functions, reinforcing the modern analytic foundations of quaternion-valued mathematics.

At the same time, practical applications of quaternionic probability and statistics have emerged in signal processing and image analysis. For instance, quaternionic Fourier transforms [25], principal component analysis [24], and augmented statistics of quaternion random variables [23] have shown how quaternion-valued frameworks can capture multi-dimensional correlations in ways that real or complex approaches cannot. Recent mathematical work, such as Wang and Zhang's quaternionic Mahler measure [26], further illustrates that the study of measures with values in $\mathbb{H}$ is both mathematically natural and actively developing.

Despite these advances, a systematic axiomatic theory of quaternionic probability—comparable in rigor to Kolmogorov's framework, but extending it into the noncommutative domain of quaternions—has not yet been established. The goal of this work is to provide such a theory: to define quaternionic probability spaces, state coherent axioms, and develop the corresponding notions of conditional probability, independence, random variables, processes, and information measures. By doing so, we show how classical probability [1–5] and quantum complex probability [6,7,9] appear as special cases, while new phenomena—such as order-dependent conditional probabilities, noncommutative independence, and quaternionic transport equations—arise uniquely in the quaternionic setting.

This paper is structured as follows. In Section 2 we introduce the axioms of quaternionic probability and establish their basic consequences. Sections 3 and 4 develop conditional probabilities, Bayes' rules, and notions of independence in the noncommutative context. In Sections 5 and 6 we define quaternionic random variables, expectations, and observables. Sections 7 and 8 extend the theory to stochastic processes and derive quaternionic continuity and Fokker–Planck-type equations. Section 9 studies quaternionic Markov chains, while Section 10 develops a framework for quaternionic information theory. In Section 11 we compare the quaternionic framework to classical and complex probability, and in Section 12 we present illustrative examples and potential applications. Technical details are collected in Section 16.

To the best of our knowledge, no complete axiomatization of probability over the quaternionic algebra has been established in the literature. Previous works have considered noncommutative probability in the context of operator algebras or complex Hilbert spaces, but a direct extension of Kolmogorov's axioms to quaternions has not been systematically developed. The present work aims to fill this gap by constructing a rigorous quaternionic probability theory and by illustrating its distinctive features through both mathematical results and concrete examples.

2. Quaternionic Probability Spaces and Axioms

We begin by extending Kolmogorov's axioms [1] to quaternion-valued measures. Throughout, let $\mathbb{H}$ denote the skew field of quaternions, with scalar part $Sc(q)$ and vector part $Vec\ (q)$.The set of unit quaternions is denoted by $\mathbb{S}^3 = \{q \in \mathbb{H}: |q| = 1\}$.

Definition 2.1 (Quaternionic probability space).

A quaternionic probability space is a triple

$$(\Omega, \mathcal{F}, \mathbb{P}_{\mathbb{H}}),$$

where:

- Ω a sample space,
- $\mathcal{F}$ is a σ -algebra of subsets of Ω,
- $\mathbb{P}_{\mathbb{H}}: \mathcal{F} \to \mathbb{H}$ is a set function satisfying the axioms below.

Axiom 2.1 (Normalization).

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

$$Sc\mathbb{P}_{\mathbb{H}}(\Omega) = 1.$$

Axiom 2.2 (Quaternionic σ -additivity).

For any countable collection of pairwise disjoint sets $E_1, E_2, \ldots \in \mathcal{F}$,

$$\mathbb{P}_{\mathbb{H}}\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mathbb{P}_{\mathbb{H}}(E_n),$$

where convergence is taken in the quaternionic norm.

Axiom 2.3 (Scalar positivity).

$$Sc\ \mathbb{P}_{\mathbb{H}}(E) \geq 0 \text{ for all } E \in \mathcal{F}.$$

Axiom 2.4 (Complement).

$$\mathbb{P}_{\mathbb{H}}(E^C) = 1 - \mathbb{P}_{\mathbb{H}}(E).$$

Axiom 2.5 (Boundedness).

$$|\mathbb{P}_{\mathbb{H}}(E)| \leq 1 \text{ for all } E \in \mathcal{F}\ .$$

Lemma 2.1 (Classical reduction).

The scalar part

$$P(E) \coloneqq Sc\ \mathbb{P}_{\mathbb{H}}(E).$$

defines a classical probability measure on $(\Omega, \mathcal{F})$.

Proof. Immediate from Axioms 2.1–2.3.

Proposition 2.1 (Modulus–phase factorization).

Every quaternionic probability can be uniquely written as

$$\mathbb{P}_{\mathbb{H}}(E) = P(E)U(E),$$

where $\mathbb{P}(E) \in [0,1]$ is the classical probability of E and $U(E) \in S^3$ is a unit quaternion.
Interpretation. The scalar component $\mathbb{P}(E)$ captures the usual frequency interpretation, while the unit quaternion $U(E)$ encodes an additional "phase" or orientation, absent in the classical case.
This axiomatic structure establishes quaternionic probability as a genuine extension of Kolmogorov's framework. In the next section we explore the consequences for conditional probabilities and Bayes' rule in a noncommutative setting.

3. Conditional Probabilities and Bayes' Theorem.

One of the main novelties of quaternionic probability is that conditionalization becomes inherently noncommutative. Because quaternionic multiplication does not commute, we must distinguish between *right* and *left* conditional probabilities.

Definition 3.1 (Right conditional probability).

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

For events $A, B \in \mathcal{F}$ with $\mathbb{P}_{\mathbb{H}}(B) \neq 0$, the right conditional probability of A given B is defined by

$$\mathbb{P}^R(A|B) \coloneqq \mathbb{P}_{\mathbb{H}}(A \cap B)\left(\mathbb{P}_{\mathbb{H}}(B)\right)^{-1}.$$

Definition 3.2 (Left conditional probability).

For events $A, B \in \mathcal{F}$ with $\mathbb{P}_{\mathbb{H}}(B) \neq 0$, the left conditional probability of A given B is defined by

$$\mathbb{P}^L(A|B) \coloneqq \left(\mathbb{P}_{\mathbb{H}}(B)\right)^{-1}\mathbb{P}_{\mathbb{H}}(A \cap B).$$

Proposition 3.1 (Scalar reduction).

For both left and right conditionals,

$$Sc\,\mathbb{P}^R(A|B) = Sc\,\mathbb{P}^L(A|B) = P(A|B).$$

where P $P(A|B)$ is the classical conditional probability.

Proof. Since scalar parts commute, the difference between left and right multiplications disappears when taking $Sc(\cdot)$.

Proposition 3.2 (Order-dependence).

In general,

$$\mathbb{P}^R(A|B) \neq \mathbb{P}^L(A|B).$$

Thus, conditional probabilities encode not only relative frequency but also the order of information update.

Theorem 3.1 (Right Bayes' rule).

Let $\{B_j\}$ be a finite or countable partition of Ω with $\mathbb{P}_{\mathbb{H}}(B_j) \neq 0$. Then, for any $A \in \mathcal{F}$,

$$\mathbb{P}^R(B_k|A) = \frac{\mathbb{P}^R(A|B_k)\,\mathbb{P}_{\mathbb{H}}(B_k)}{\sum_j \mathbb{P}^R(A|B_j)\,\mathbb{P}_{\mathbb{H}}(B_j)}.$$

Theorem 3.2 (Left Bayes' rule).

Under the same assumptions,

$$\mathbb{P}^L(B_k|A) = \left(\sum_j \mathbb{P}^R(A|B_j)\,\mathbb{P}_{\mathbb{H}}(B_j)\right)^{-1}\mathbb{P}^R(A|B_k)\,\mathbb{P}_{\mathbb{H}}(B_k).$$

Remark 3.1.

- Both Bayes' rules reduce to the classical one when all quaternionic phases commute.
- The distinction between left and right formulations reflects the noncommutative geometry of the quaternionic unit sphere $\mathbb{S}^3$.
- This framework naturally accommodates order-sensitive inference, where updating on A then B is not equivalent to updating on B then A.

This section demonstrates that conditionalization in quaternionic probability is fundamentally noncommutative, giving rise to two inequivalent but consistent Bayes' rules. In the next section we will investigate how these asymmetries manifest in different notions of independence.

4. Independence and Correlations.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

In the quaternionic framework, the noncommutativity of multiplication implies that the notion of independence is no longer unique. We distinguish several types of independence and define quaternionic correlation.

Definition 4.1 (Strong independence).

Two events $A, B \in \mathcal{F}$ are said to be strongly independent if

$$\mathbb{P}_{\mathbb{H}}(A \cap B) = \mathbb{P}_{\mathbb{H}}(A)\, \mathbb{P}_{\mathbb{H}}(B).$$

Definition 4.2 (Right independence).

Events A, B are right independent if

$$\mathbb{P}_{\mathbb{H}}(A \cap B) = \mathbb{P}^{R}(A|B)\, \mathbb{P}_{\mathbb{H}}(B).$$

Definition 4.3 (Left independence).

Events A, B are left independent if

$$\mathbb{P}_{\mathbb{H}}(A \cap B) = \mathbb{P}_{\mathbb{H}}(B) \mathbb{P}^{L}(A|B).$$

Definition 4.4 (Scalar independence).

Events A, B are scalar independent if

$$P(A \cap B) = P(A)P(B),$$

where $P = Sc\, \mathbb{P}_{\mathbb{H}}$ is the underlying classical probability.

Proposition 4.1.

Strong independence $\Longrightarrow$ right independence, left independence, and scalar independence. The converses do not hold in general.

Proof. Strong independence directly implies equality under both left and right conditional definitions and yields the scalar condition. However, the reverse implications fail when quaternionic phases do not commute.

Definition 4.5 (Quaternionic correlation).

The quaternionic correlation between two events A, B is defined as

$$C_{\mathbb{H}}(A, B) \coloneqq \mathbb{P}_{\mathbb{H}}(A \cap B) - \mathbb{P}_{\mathbb{H}}(A)\, \mathbb{P}_{\mathbb{H}}(B).$$

Proposition 4.2 (Scalar reduction of correlation).

The scalar part of quaternionic correlation is the classical covariance:

$$Sc\, C_{\mathbb{H}}(A, B) \coloneqq P(A \cap B) - P(A)P(B).$$

Remark 4.1.

- In the classical case, independence is unique; here, different forms of independence reflect the underlying noncommutative structure.
- The quaternionic correlation captures deviations not only in scalar probability but also in vectorial phase coherence.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

- Order effects (left vs. right independence) provide a natural mathematical description of phenomena where the sequence of conditioning matters.

This section shows that the richness of quaternionic probability emerges already at the level of independence. Multiple notions coexist, and correlations now carry both scalar and vectorial information. In the next section we will introduce quaternionic random variables and define expectations and moments.

5. Quaternionic Random Variables and Expectation.

Having defined quaternionic probability measures, we now extend the framework to random variables, expectations, and moments. The key novelty is that expectations become quaternion-valued, and inequalities must be interpreted via their scalar components.

Definition 5.1 (Quaternionic random variable).

A quaternionic random variable is a measurable map

$$X: \Omega \to \mathbb{H}.$$

The distribution of X is the quaternionic measure

$$\mu_X(B) := \mathbb{P}_{\mathbb{H}}(X^{-1}(B)), \qquad B \subseteq \mathbb{H}.$$

Definition 5.2 (Expectation).

The expectation of X is defined by

$$\mathbb{E}_{\mathbb{H}}[X] = \int_{\Omega} X(\omega) d\mathbb{P}_{\mathbb{H}}(\omega).$$

When $\mathbb{P}_{\mathbb{H}}$ is expressed in modulus–phase form $\mathbb{P}_{\mathbb{H}} = PU$, this integral can be interpreted as

$$\mathbb{E}_{\mathbb{H}}[X] = \int_{\Omega} X(\omega) U(\omega) dP(\omega)$$

with P the scalar probability and $U(\omega) \in \mathbb{S}^3$ a local quaternionic phase.

Proposition 5.1 (Scalar expectation).

The scalar part of the quaternionic expectation is the classical expectation:

$$Sc\, \mathbb{E}_{\mathbb{H}}[X] = \mathbb{E}[X],$$

where $\mathbb{E}[X]$ is the usual expectation with respect to $P = Sc\, \mathbb{P}_{\mathbb{H}}$.

Definition 5.3 (Variance).

The variance of X is defined as

$$Var_{\mathbb{H}}(X) := \mathbb{E}_{\mathbb{H}}[|X|^2] - |\mathbb{E}_{\mathbb{H}}[X]|^2$$

Proposition 5.2 (Scalar variance).

The scalar part of the quaternionic variance coincides with the classical variance:

$$Sc\, Var_{\mathbb{H}}(X) = Var(X).$$

Theorem 5.1 (Markov inequality, quaternionic form).

Let $X \geq 0$ be a real-valued random variable on a quaternionic probability space. Then, for any $\alpha > 0$,

$$P(X \geq \alpha) \leq \frac{\mathbb{E}[X]}{\alpha}$$

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

where $P = Sc\,\mathbb{P}_{\mathbb{H}}$.

Remark. The inequality remains scalar because order in $\mathbb{H}$ is not defined; only the real measure governs tail bounds.

Theorem 5.2 (Chebyshev inequality, quaternionic form).

For any real-valued X with finite variance,

$$P(|X - \mathbb{E}[X]| \geq \varepsilon) \leq \frac{\mathbb{E}[X]}{\varepsilon^2}$$

Again, the scalar distribution P controls concentration, while the quaternionic components contribute through higher-order coherence rather than tail probabilities.

Remark 5.1.

- Expectations and variances carry quaternionic information, but inequalities remain scalar because order is defined only in $\mathbb{R}$.
- The vector components of $\mathbb{E}[X]$ may encode directional or phase-like features of the distribution.

This section establishes the basic theory of quaternionic random variables. In the next section we introduce projections onto complex subalgebras and observables, showing how real measurements can be extracted from quaternionic distributions.

6. Transformations, Projections, and Observables.

While quaternionic random variables may take values in the full noncommutative algebra $\mathbb{H}$, in practice one often measures or analyzes their behavior through projections onto real or complex subalgebras. This section formalizes such transformations and introduces the notion of observables.

Definition 6.1 (Projection onto a subalgebra).

For a fixed imaginary unit $u \in \{i, j, k\}$, define the projection

$$\pi_u : \mathbb{H} \to \mathbb{C}_u := \{a + bu : a, b \in \mathbb{R}\}$$

by

$$\pi_u(a + bi + cj + dk) = a + bu.$$

Proposition 6.1.

For any quaternionic random variable X, the projection $X^{(u)} := \pi_u(X)$ is a complex-valued random variable with respect to the underlying scalar probability measure $P = Sc\,\mathbb{P}_{\mathbb{H}}$.

Definition 6.2 (Observable).

Let $X : \Omega \to \mathbb{H}$ be a quaternionic random variable and $u \in \{i, j, k\}$. The observable associated with direction uuu is defined as

$$O_u(X) := Sc\,(X\,u^*).$$

This extracts a real-valued random variable aligned with the chosen imaginary axis.

Proposition 6.2 (Expectation of observables).

For each uuu, the expectation of the observable satisfies

$$\mathbb{E}[O_u(X)] = Sc\,\mathbb{E}_{\mathbb{H}}[X\,u^*].$$

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

Theorem 6.1 (Reduction to classical/complex cases).

- If $\mathrm{U(E)} \equiv 1$ for all E, then quaternionic probability reduces to classical probability.
- If all phases $\mathrm{U(E)}$ lie within a single complex subalgebra $\mathbb{C}_{\mathrm{u}} \subset \mathbb{H}$, then quaternionic probability reduces to complex probability, structurally equivalent to standard quantum mechanics.

Remark 6.1.

- Projections and observables provide the link between quaternionic models and experimentally accessible quantities.
- This mirrors the role of self-adjoint operators in quantum mechanics, but extended to the richer quaternionic setting.

This section clarifies how quaternionic random variables can be connected to measurable outputs. In the next section we will extend the framework to stochastic processes and derive quaternionic versions of balance and continuity equations.

7. Quaternionic Processes and Continuity Equations.

In classical probability, stochastic processes are families of random variables indexed by time, whose evolution is described by balance or continuity equations. In the quaternionic setting, these structures generalize to quaternion-valued densities and currents.

Definition 7.1 (Quaternionic stochastic process).

A **quaternionic stochastic process** is a family of quaternionic random variables

$$\{X_t t: \Omega \to \mathbb{H}\}_{t \geq 0},$$

adapted to some filtration $\{\mathcal{F}_t\}_{t \geq 0}$.

Definition 7.2 (Quaternionic probability density).

A quaternionic-valued density is a function

$$\rho(x,t) = \rho_0(x,t) + \rho_1(x,t)i + \rho_2(x,t)j + \rho_3(x,t)k$$

with $\rho_\ell(x,t) \in \mathbb{R}$. Its scalar part ρ_0 represents the classical probability density.

Definition 7.3 (Quaternionic current).

The **probability current** is a quaternionic vector field

$$J(x,t) = J_0(x,t) + J_1(x,t)i + J_2(x,t)j + J_3(x,t)k$$

with $J_m(x,t) \in \mathbb{R}^d$.

Theorem 7.1 (Quaternionic continuity equation).

For any region $V \subset \mathbb{R}^d$,

$$\frac{d}{dt}\int_V \rho(x,t)dx = -\int_{\partial V} J(x,t) \cdot n \, dS + \int_V S(x,t)dx,$$

where $S(x,t)$ is a quaternionic source term.

In differential form,

$$\partial_t \rho(x,t) + \nabla \cdot J(x,t) = S(x,t).$$

Proof. Follows from quaternionic integration and the divergence theorem applied componentwise.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

Proposition 7.1 (Scalar reduction).

The scalar part of the continuity equation is

$$\partial_t \rho_0(x,t) + \nabla \cdot J_0(x,t) = S_0(x,t).$$

which is the classical continuity equation.

Remark 7.1.

- The scalar component governs ordinary conservation of probability.
- The vectorial components describe additional degrees of freedom, such as orientation or phase transport in quaternionic space.
- This structure allows for richer dynamics while ensuring that the real part remains a valid probability density.

This section establishes the quaternionic continuity equation as the fundamental balance law of quaternionic probability. In the next section we derive transport equations and diffusion dynamics, including a quaternionic Fokker–Planck equation.

8. Transport and the One-Dimensional Case.

The quaternionic continuity equation provides the general balance law for probability densities. In this section, we introduce specific transport models and study their one-dimensional reduction.

Definition 8.1 (Quaternionic drift–diffusion current).

Given a velocity field $v \in \mathbb{H}^d$ and diffusion coefficient $D > 0$, the quaternionic probability current is modeled as

$$J(x,t) = v\,\rho(x,t) - D\nabla\rho(x,t).$$

Theorem 8.1 (Quaternionic Fokker–Planck equation).

Substituting the drift–diffusion current into the continuity equation yields

$$\partial_t \rho(x,t) = -(v \cdot \nabla)\rho(x,t) + D\nabla^2\rho(x,t).$$

Proof. Direct substitution and application of the divergence theorem.

Proposition 8.1 (Component form).

Writing $\rho = \rho_0 + \rho_1 i + \rho_2 j + \rho_3 k$, each component satisfies

$$\partial_t \rho_m(x,t) = -(v \cdot \nabla)\rho_m(x,t) + D\nabla^2\rho_m(x,t), \qquad m = 0,1,2,3.$$

Corollary 8.1 (Scalar reduction).

The scalar part ρ_0 evolves according to the classical Fokker–Planck equation, while the vector parts ρ_1, ρ_2, ρ_3 evolve as coupled drift–diffusion fields without affecting normalization.

Definition 8.2 (One-dimensional transport).

In one dimension,

$$\partial_t \rho_m(x,t) = -\partial_x\big(v\rho_x(x,t)\big) + D\partial_x^2\rho_x(x,t).$$

Proposition 8.2 (Solution structure in 1D).

For constant v and D, the solution takes the form

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

$$\rho(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{(x-vt)^2}{4Dt}\right) U,$$

where $U \in \mathbb{S}^3$ is a unit quaternion encoding the initial phase orientation.

Remark 8.1.

- The one-dimensional case illustrates how quaternionic probability preserves the scalar Gaussian kernel while enriching it with quaternionic phase structure.
- The evolution of vectorial components can be interpreted as rotational diffusion in quaternionic space.

This section demonstrates how transport and diffusion extend naturally into the quaternionic setting, reducing to classical equations in the scalar component but enriching the dynamics with quaternionic phases. In the next section, we introduce quaternionic Markov chains and their generators.

9. Quaternionic Markov Chains and Generators

Markov processes describe systems evolving by local transitions between states. In the quaternionic setting, transition probabilities are replaced by quaternionic-valued entries, introducing order-dependence and noncommutativity.

Definition 9.1 (Quaternionic stochastic matrix).

A matrix $P = (p_{ij}) \in M_n(\mathbb{H})$ is a quaternionic stochastic matrix if:

1. $p_{ij} \in \mathbb{H}$,
2. $Sc(p_{ij}) \geq 0$ for all i, j,
3. $\sum_j Sc(p_{ij}) = 1$for each row i,
4. $|p_{ij}| \leq 1$.

Definition 9.2 (Quaternionic Markov chain).

A process $\{X_t\}_{t\geq 0}$ with state space $\{1, \dots, n\}$ is a quaternionic Markov chain if

$$\mathbb{P}_{\mathbb{H}}(X_{t+1} = j \mid X_t = i, history) = p_{ij},$$

with (p_{ij}) a quaternionic stochastic matrix.

Proposition 9.1 (Evolution of distributions).

If $\pi^{(0)} \in \mathbb{H}^n$ is the initial distribution, then

$$\pi^{(t+1)} = \pi^{(t)} P$$

Lemma 9.1 (Scalar reduction).

The scalar part of $\pi^{(t)}$ evolves as a classical Markov chain with transition matrix $Sc(P)$.

Definition 9.3 (Quaternionic generator).

In continuous time, a quaternionic generator is a matrix $Q = (q_{ij}) \in M_n(\mathbb{H})$ satisfying:

5. $Sc(q_{ij}) \geq 0$ for $i \neq j$.
6. $\sum_j Sc(q_{ij}) = 0$ for each row i.

The evolution is given by

$$\pi(t) = \pi(0) e^{Qt}$$

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

Theorem 9.1 (Semigroup property).

The family $T(t) = e^{Qt}$ forms a semigroup preserving scalar normalization:

$$\sum_i Sc(\pi_i(t)) = 1 \quad \text{for all } t \geq 0.$$

Definition 9.4 (Stationary distribution).

A vector $\pi \in H^n$ is **stationary** if

$$\pi P = \pi \ (discrete\ time), \qquad \pi Q = 0 \ (continuous\ time).$$

Theorem 9.2 (Scalar stationarity).

If π is stationary in the quaternionic sense, then $Sc(\pi)$ is a stationary distribution for the associated classical Markov chain.

Remark 9.1.

- The scalar structure ensures consistency with classical Markov theory.
- Vectorial components enrich the model by introducing noncommutative order effects and quaternionic phase rotations.
- Applications include systems where both transition frequencies and phase-like correlations are relevant.

This section establishes quaternionic Markov processes, extending the classical theory with quaternion-valued transition structures. In the next section, we build on this framework to define entropy, divergence, and information measures in quaternionic probability.

10. Quaternionic Information Theory

Information theory provides a quantitative framework to study uncertainty, divergence, and coherence in probability distributions. Extending these notions to quaternionic probability requires separating scalar uncertainty from vectorial coherence.

Definition 10.1 (Scalar entropy).

For a quaternionic density $\rho(x)$, the scalar entropy is defined as

$$H(\rho_0) = -\int \rho_0(x) \log \rho_0(x)\, dx,$$

where $\rho_0 = Sc(\rho)$.

This coincides with the classical Shannon entropy.

Definition 10.2 (Vector coherence functional).

The vectorial coherence of ρ is defined as

$$I_{vec}(\rho) = \int \|Vec(\rho(x))\| dx.$$

- $I_{vec} = 0$: purely classical distribution.
- $I_{vec} > 0$: quaternionic coherence beyond real probability.

Definition 10.3 (Quaternionic entropy).

The quaternionic entropy is defined by

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

$$H_{\mathbb{H}}(\rho) = H(\rho_0) + \lambda I_{vec}(\rho)$$

where $\lambda > 0$ is a weighting parameter.

Definition 10.4 (Quaternionic divergence).

For two densities ρ,σ, define

$$D_{\mathbb{H}}(\rho \parallel \sigma) = \int [\rho_0(x) \log \frac{\rho_0(x)}{\sigma_0(x)} + \left\| Vec(\rho(x)) - Vec(\sigma(x)) \right\|^2] dx.$$

- The first term is the classical Kullback–Leibler divergence.
- The second measures the mismatch of quaternionic vector phases.

Theorem 10.1 (Non-negativity).

$$D_{\mathbb{H}}(\rho \parallel \sigma) \geq 0, \qquad and\ D_{\mathbb{H}}(\rho \parallel \sigma) \Leftrightarrow \rho = \sigma$$

Theorem 10.2 (Monotonicity under channels).

If Φ is a linear map preserving scalar normalization, then

$$D_{\mathbb{H}}(\Phi\ (\rho) \parallel \Phi\ (\sigma)) \leq D_{\mathbb{H}}(\rho \parallel \sigma)$$

This extends the classical data-processing inequality.

Remark 10.1.

- The scalar entropy reflects classical uncertainty.
- The coherence functional captures quaternionic information not visible in real probabilities.
- Divergence in quaternionic probability simultaneously measures probabilistic and orientational mismatch.

This section introduces entropy, coherence, and divergence in quaternionic probability, laying the foundation for a broader information-theoretic framework. In the next section, we compare the quaternionic model with its classical and complex counterparts.

11. Comparison with Classical and Complex Cases.

The quaternionic framework simultaneously generalizes both classical Kolmogorov probability and complex quantum probability. This section highlights how these theories are embedded as special cases.

Proposition 11.1 (Classical reduction).

If all quaternionic phases are trivial, $U(E) \equiv 1$, then

$$\mathbb{P}_{\mathbb{H}}(\mathrm{E}) = \mathbb{P}(E) \in [0,1] \subset \mathbb{R}.$$

- Left and right conditional probabilities coincide.
- Independence reduces to the classical definition.
- The continuity and transport equations reduce to their classical forms.

Thus, classical probability [1–5] is a subtheory of quaternionic probability.

Proposition 11.2 (Complex reduction).

If all phases $U(E)$ lie in a fixed complex subalgebra $\mathbb{C}_u \subset \mathbb{H}$, then

$$\mathbb{P}_{\mathbb{H}}(\mathrm{E}) \in \mathbb{C}_u$$

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

- The theory reduces to complex probability, equivalent to quantum mechanics with amplitudes in $U(1)$.
- Conditional probabilities and Bayes' rule become the complex-valued forms familiar in quantum inference.

Proposition 11.3 (Strict quaternionic phenomena).

There exist effects with no classical or complex analogue:

1. Order dependence in conditionals:

$$\mathbb{P}^R(A|B) \neq \mathbb{P}^L(A|B),$$

even when scalar probabilities coincide.

2. Noncommutative independence:

Strong, left, right, and scalar independence diverge in meaning.

3. Quaternionic phase rotations:

Phases evolve in $SU(2) \simeq \mathbb{S}^3$, a non-Abelian group, unlike the Abelian $U(1)$ of complex quantum theory.

Remark 11.1.

- The real case corresponds to classical probability.
- The complex case corresponds to standard quantum mechanics.
- The quaternionic case yields a genuine noncommutative probability theory, where the order of updates, correlations, and transport laws are fundamentally enriched.

12. Examples and Applications.

Example 12.1 — Order dependence in medical diagnostics.

Let $A =$"test positive for marker 1" and $B =$ "test positive for marker 2", with

$$P_{\mathbb{H}}(A) = \frac{1}{2}(1+i), \;\; P_{\mathbb{H}}(B) = \frac{1}{2}(1+j)$$

Then

$$P_{\mathbb{H}}(A|B)_{left} \neq P_{\mathbb{H}}(A|B)_{right},$$

although both have scalar part 1/2.

As shown in Figure 1, the scalar components are equal, but the imaginary parts differ depending on conditioning order. This models order-sensitive diagnostic protocols, where the sequence of tests changes the inferred probability.

Example 12.2 — Quaternionic correlation in orthogonal signals.

Consider signals $X = sin\,(t)$ and $Y = cos\,(t)$. Classically,

$$Cov(X,Y) = 0,$$

so they appear independent.

However, in the quaternionic framework, X and Y form a circularly polarized pair: their joint trajectory lies on a circle in the sine–cosine plane. This reveals perfect coherence despite zero covariance.

As illustrated in Figure 2, the classical view treats them as independent signals, while the quaternionic view identifies a circular correlation structure.

Example 12.3 — Quaternionic diffusion in spintronics.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

The 1D diffusion equation gives

$$\rho(x,t) = \frac{1}{\sqrt{4\pi Dt}} e^{-x^2} \cdot u(t), \qquad u(t) = e^{jt} i$$

The scalar density follows the Gaussian law, while $u(t)$ rotates over time.

In Figure 3, the Gaussian profiles broaden with increasing t, while the quaternionic phase evolves, capturing joint density–spin dynamics relevant for spin-polarized currents.

Example 12.4 — Market cycles with quaternionic Markov chains.

For the transition matrix

$$M = \begin{pmatrix} 0 & 1 \\ e^{i\theta} & 0 \end{pmatrix},$$

the scalar probabilities alternate deterministically (bull–bear).

In Figure 4, the classical chain alternates states, while the quaternionic version introduces a rotating phase. This models cyclical market behavior with memory, adding structure beyond frequency counts.

Example 12.5 — Divergence of orthogonal distributions.

Consider two distributions with the same scalar Gaussian density but orthogonal quaternionic orientations:

$$P(x) = \rho(x) \cdot i, \qquad Q(x) = \rho(x) \cdot j$$

Classically,

$$D_{KL}(P \parallel Q) = 0,$$

so they are indistinguishable. Quaternionically,

$$D_{KL}(P \parallel Q) > 0,$$

since i and j are orthogonal.

As shown in Figure 5, the scalar view sees identical Gaussians, while the quaternionic view separates them as orthogonal distributions in phase space. This demonstrates the ability of quaternionic divergence to detect differences invisible to classical statistics.

13. Conclusion and Outlook.

This work introduces an axiomatic framework for quaternionic probability, extending Kolmogorov's real-valued theory to measures and expectations taking values in $\mathbb{H}$ .The noncommutativity of quaternions manifests in two key ways: (i) left/right conditional probabilities and order-sensitive Bayes' rules, and (ii) multiple, non-equivalent notions of independence (strong, left, right, scalar). We developed the associated calculus of random variables, expectation and variance, and we derived continuity and Fokker–Planck-type equations for quaternionic densities and currents. The framework accommodates quaternionic Markov chains and proposes a natural information theory separating scalar uncertainty from vectorial coherence.

The scalar projection recovers classical probability, while restriction to a complex subalgebra recovers complex (quantum-like) probability; yet several phenomena are strictly quaternionic, including order-dependent conditioning and phase-rotational transport in $SU(2)$. The examples show concrete scenarios where classical statistics is blind (identical scalars, zero KL) but quaternionic predictions differ in observables, divergences, or dynamics.

Future directions include: (1) law of large numbers and central limit theorems with vectorial coherence; (2) martingales, filtrations, and Doob-type results in quaternionic settings; (3) ergodic theory and spectral gaps for quaternionic Markov generators; (4) inference algorithms and estimation of quaternionic phases from data; (5) applications to spin-polarized transport, orientation-aware sensing, and order-effects in decision pipelines.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

The examples presented in Section 12 highlight practical scenarios where quaternionic probability provides information that classical and even complex-valued frameworks cannot capture. In medical diagnostics, it naturally models order effects in sequential tests; in signal analysis, it reveals hidden coherence between orthogonal channels; in spintronics, it describes coupled diffusion and spin precession within a single law; in financial Markov chains, it encodes cyclical memory beyond scalar transitions; and in communications, it distinguishes channels with identical power spectra but different polarization states. These applications show that quaternionic probability is not only a consistent mathematical generalization but also a tool with concrete advantages in modeling systems where orientation, coherence, and order-dependence play a fundamental role.

Future directions include the exploration of connections between quaternionic probability and operator-algebraic approaches to noncommutative probability, its potential role in quaternionic quantum mechanics, and possible links with free probability and noncommutative geometry. These perspectives suggest that the present axiomatization may serve not only as a self-contained framework but also as a starting point for broader developments in mathematics and physics.

14. Numerical Methods and Reproducibility.

All figures were produced with Python and Matplotlib (no seaborn, one plot per figure, default color cycle). The code is deterministic and uses only elementary numerical routines. Source scripts are available upon request and can be deposited in a public repository.

Computational environment.

- Language: Python 3.x; plotting: Matplotlib.
- No stochastic seeds are required; outputs are deterministic.
- Hardware: standard laptop; no GPU or special libraries.

15. Figure captions.

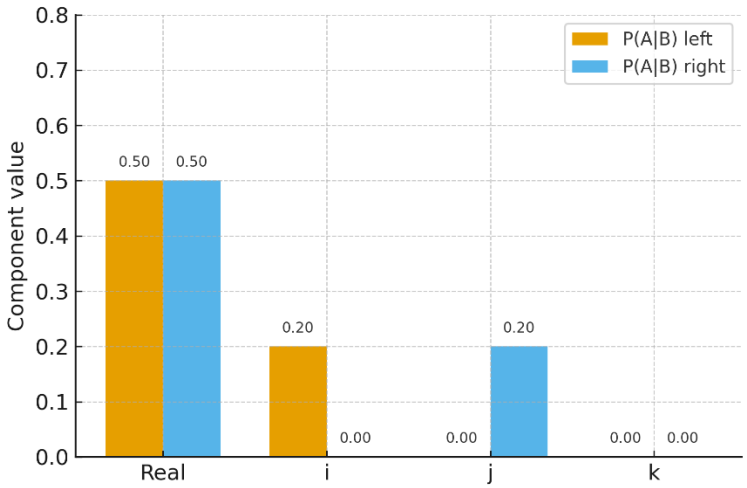

Figure 1. Left vs. right conditional probabilities. Both share the same scalar part but differ in their imaginary components, capturing order-dependent effects relevant in diagnostic testing.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

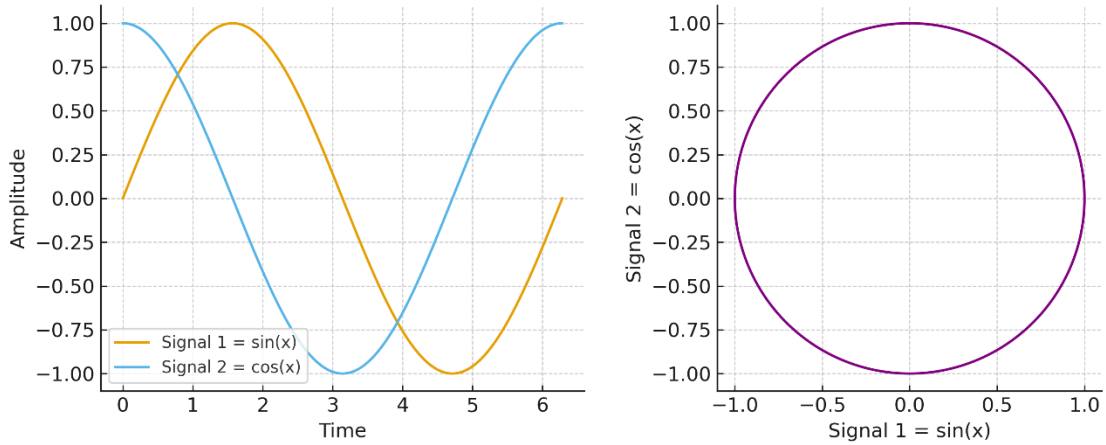

Figure 2. Classical vs. quaternionic correlation of sine and cosine signals. Classical covariance is zero (independence), while the quaternionic representation reveals perfect circular coherence.

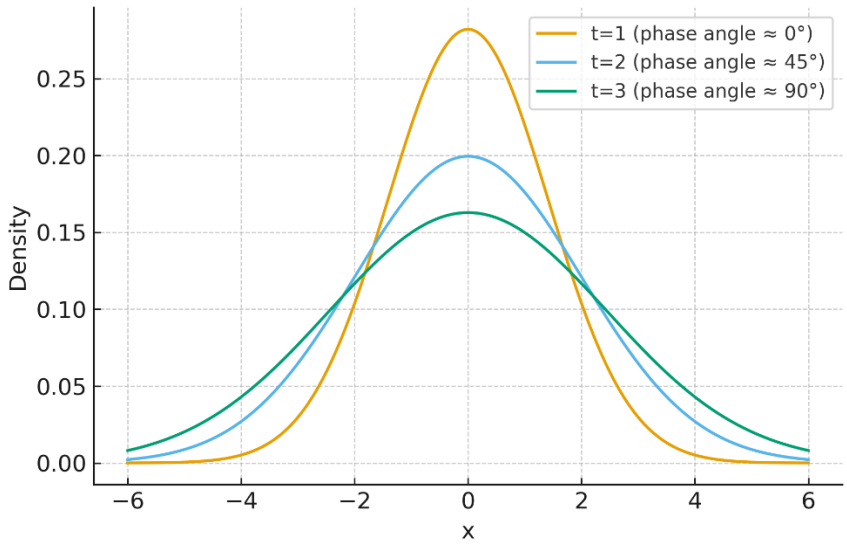

Figure 3. Scalar Gaussian diffusion at times $t = 1,2,3$ The scalar density broadens with time, while the quaternionic phase evolves, modeling coupled density–spin dynamics.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

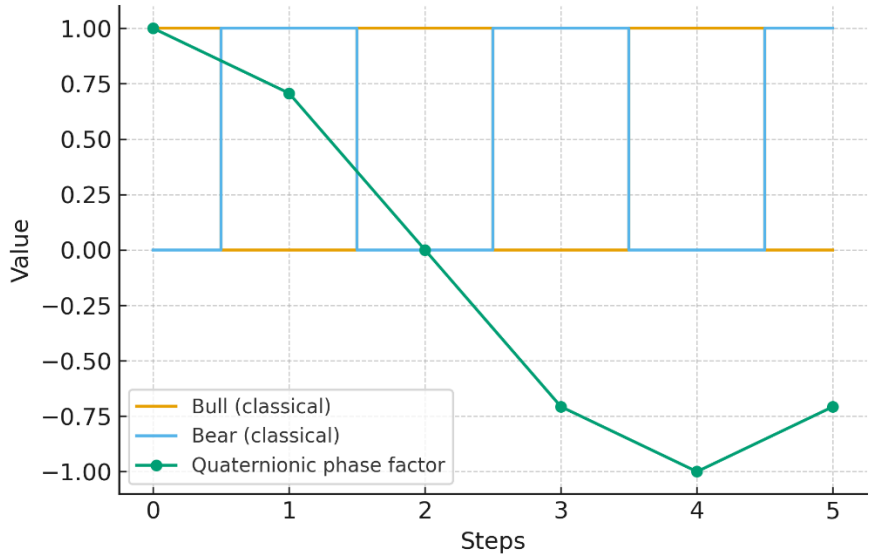

Figure 4. Two-state Markov chain. Classical probabilities alternate deterministically between states, while quaternionic evolution introduces a rotating phase, modeling cyclical memory effects.

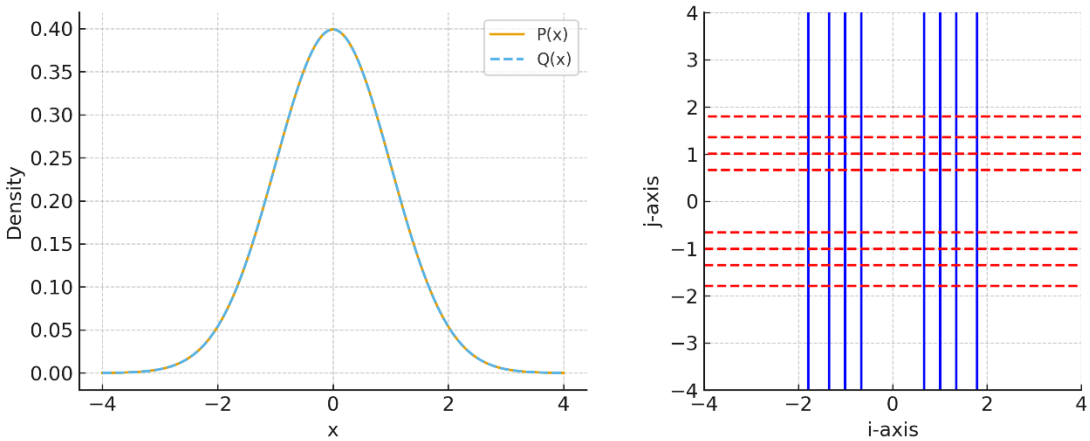

Figure 5. Scalar view: two identical Gaussian distributions, indistinguishable with classical KL divergence ($D_{KL} = 0$). Quaternionic view: orthogonal orientations in phase space, yielding positive quaternionic divergence ($D_{\mathbb{H}} > 0$).

16. Technical Appendices

This appendix collects mathematical foundations and technical proofs that support the quaternionic probability framework.

16.1. Quaternionic measures

Definition A.1 (Quaternionic measure).

A function $\mu: \mathcal{F} \to \mathbb{H}$ is a quaternionic measure if:

1. $\mu(\emptyset) = 0$.
2. For pairwise disjoint $E_1, E_2, \cdots \in \mathcal{F}$,

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n),$$

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

with convergence in the quaternionic norm.

Lemma A.1 (Total variation).

Define the total variation of μ by

$$|\mu|(E) = \sup\left\{\sum_j |\mu(E_j) : \{E_j\}\, finite\ partition\ of\ E\right\}.$$

Then $|\mu|$ is a real-valued measure and μ is absolutely continuous with respect to $|\mu|$.

16.2. Quaternionic integration.

Definition A.2 (Bochner–quaternionic integral).

For a simple function $f = \sum_k q_k \mu(E_k)$, define

$$\int f d\mu = \sum_k q_k \mu(E_k)$$

This extends by norm limits to all measurable functions $f: \Omega \to \mathbb{H}$.

Proposition A.1 (Norm inequality).

For any $f \in L^{-1}(\mu)$,

$$\left|\int f d\mu\right| \leq \int |f|\, d|\mu|$$

16.3 Functional calculus in $\mathbb{H}$.

- Exponential:

$$e^q = \sum_{n=0}^{\infty} \frac{q^n}{n!}, \quad q \in \mathbb{H}.$$

- Logarithm: defined on $\mathbb{H} \setminus (-\infty, 0]$, with multiple values corresponding to directions in $\mathbb{S}^3$.
- S-spectrum: For operators on quaternionic Banach spaces, the S-spectrum (Colombo–Sabadini–Struppa [22]) generalizes the classical spectrum and is essential for defining e^{Qt} in Section 9.

16.4. Proof sketches.

A.4.1 Bayes' rules.

Noncommutativity requires distinguishing left and right multiplication. The denominator normalizes the scalar part to unity, ensuring consistency with Definition 3.1–3.2.

A.4.2 Continuity equation.

Derived by integrating the balance law over regions and applying Gauss' theorem componentwise.

A.4.3 Fokker–Planck equation.

Obtained as a diffusion limit of jump processes with quaternionic increments, scaling $\Delta x^2 \sim \Delta t$.

16.5. Summary.

- Quaternionic measures extend real measures consistently.
- The Bochner integral provides a rigorous foundation for expectations.
- Functional calculus guarantees the well-definedness of semigroups e^{Qt}.
- Proofs of conditional probability, continuity, and transport equations ensure mathematical coherence of the quaternionic framework.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 September 2025 doi:10.20944/preprints202509.2314.v1

This appendix secures the mathematical foundations of quaternionic probability, completing the proposed theory.

Data Availability Statement: This article is based entirely on mathematical derivations and theoretical reasoning. Therefore, no datasets were generated or analyzed, and data sharing is not applicable to this article.

References

1. Kolmogorov, A. N. (1933). *Foundations of the theory of probability*. Chelsea Publishing.
2. Billingsley, P. (1995). *Probability and measure* (3rd ed.). Wiley.
3. Doob, J. L. (1953). *Stochastic processes*. Wiley.
4. Feller, W. (1968). *An introduction to probability theory and its applications* (Vol. 1, 3rd ed.). Wiley.
5. Parthasarathy, K. R. (1967). *Probability measures on metric spaces*. Academic Press.
6. Feynman, R. P. (1987). Negative probability. In B. J. Hiley & F. D. Peat (Eds.), *Quantum implications: Essays in honour of David Bohm* (pp. 235–248). Routledge.
7. Khrennikov, A. (2009). *Interpretations of probability*. De Gruyter.
8. Khrennikov, A. (1999). Classical and quantum mechanics on p-adic trees of ideas. *Doklady Mathematics, 59*(3), 450–452.
9. Accardi, L. (1984). Non-Kolmogorovian probabilities and quantum physics. In *Proceedings of the Conference on Mathematical Problems in Theoretical Physics* (pp. 147–159). Springer.
10. Accardi, L., & Cecchini, C. (1982). Conditional expectations in von Neumann algebras and a theorem of Takesaki. *Journal of Functional Analysis, 45*(2), 245–273.
11. Voiculescu, D. V. (1992). Free probability theory: Random matrices and von Neumann algebras. *Proceedings of the International Congress of Mathematicians* (pp. 227–241).
12. Speicher, R. (2007). *Lectures on free probability theory*. Springer.
13. Wysoczański, J. (2010). Non-commutative independence of algebras and applications to probability. *Operators and Matrices, 4*(2), 253–270.
14. Kümmerer, B. (1985). Markov dilations and noncommutative processes. In *Probability measures on groups* (pp. 216–236). Springer.
15. Hamilton, W. R. (1853). *Lectures on quaternions*. Hodges and Smith.
16. Sudbery, A. (1986). Quaternionic analysis. *Mathematical Proceedings of the Cambridge Philosophical Society, 85*(2), 199–225.
17. Adler, S. L. (1995). *Quaternionic quantum mechanics and quantum field*. Oxford University Press.
18. Adler, S. L. (1996). Quaternionic quantum mechanics and noncommutative dynamics. *arXiv preprint hep-th/9607008*. https://arxiv.org/abs/hep-th/9607008

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

© 2025 by the author(s). Distributed under a Creative Commons CC BY license.