# On sequential ideal measurements in the exceptional Jordan algebra

Robert R. MacGregor

hexagon:2611.00620v1

**Abstract**

Real, complex and quaternionic quantum theory have one exceptional relative, the Albert algebra $H_3(\mathbb{O})$, which single-system reconstructions leave standing. Sequences of measurements on a single system can still tell it apart, provided the measurements are ideal. A yes–no test is ideal if it disturbs the state no more than its answer requires, as under the projection postulate. Ideal tests of finitely many pure states of $H_3(\mathbb{O})$ are then matched by ideal measurements on a real, complex or quaternionic Hilbert space of any dimension exactly when those states fail to generate the algebra. When they generate it, an ideal Hilbert-space model matching every sequence from even one starting state would yield a faithful representation of the algebra and its Jordan product by operators, which Albert ruled out in 1934. Three states never generate it; whether four do is decided by one number, proportional to a squared octonion associator. Even when four do, the difference is slow to show: from the maximally mixed state, a random choice between two quaternionic qutrits matches every sequence of up to six outcomes, and only at seven is every Hilbert-space model excluded. If the tests need not be ideal, then for any finite set of tests a finite-dimensional quantum device with sharp answers matches every sequence, of any length.

**AI Statement**

All text, mathematics and proofs herein, and all Lean 4 code apart from the adapted files noted in Appendix G, were produced by Claude Fable 5.2, Claude Opus 5.5, GPT-6.2 Astra, GPT-6.2 Sol, Gemini Carbon and Gemini Hydrogen, and the main theorems on the Albert algebra are machine-checked in Lean. *Except for this statement, all content herein was generated by AI and has not been understood without AI interpretation.*

Text CC BY 4.0 Lean code Apache-2.0
Lean repository: `github.com/macgregor-sudo/albert-sequential-measurements`
ORCID: 0009-0003-1734-8054

# Contents


**1 Introduction** **1**
- 1.1 Four buttons and a light 1
- 1.2 A warm-up with a qubit and a qutrit 2
- 1.3 The exception that single-system axioms leave standing 3
- 1.4 Results 4
- 1.5 Relation to earlier work 6

**2 Sketch of the proofs** **7**

**3 Ideal tests and their sequences** **9**
- 3.1 Euclidean Jordan algebras in brief 9
- 3.2 Ideal tests 11
- 3.3 Words and their probabilities 12
- 3.4 Three model classes 13

**4 A sequential criterion for speciality** **15**
- 4.1 Statement for the Albert algebra 15
- 4.2 The mechanism: null vectors on every history 16
- 4.3 Reconstruction of a Jordan homomorphism 18
- 4.4 The converse: discard, embed, construct 20
- 4.5 General Euclidean Jordan algebras 21
- 4.6 Consequences 22

**5 Three atoms fit a qutrit; four need not** **24**
- 5.1 Three atoms always sit in a qutrit 24
- 5.2 Thirteen traces 25
- 5.3 The four-point invariant 27
- 5.4 The tetrahedron identity 30
- 5.5 How large can the invariant be? 34
- 5.6 An invariant of atoms is not a test of a table 35

**6 Finite certificates** **36**
- 6.1 Finite lists 38
- 6.2 Words of length seven suffice and length six does not 39
- 6.3 Words of length eleven for every configuration 53
- 6.4 A positive tetrahedron statistic is not a witness 55
- 6.5 No linear equality at short length 56
- 6.6 Robustness 58

**7 Five tests, every four of them quantum** **62**
- 7.1 Coordinates 62
- 7.2 A five-atom configuration 63
- 7.3 Five tests always suffice 64

**8 Where the projection postulate enters** **65**
- 8.1 A finite realization with non-Lüders instruments 65
- 8.2 Sharp effects with arbitrary post-measurement channels 67
- 8.3 Complete measurements see only transition probabilities 71

**9 Related work** **72**
- 9.1 Reconstructions, composites and embeddings 72
- 9.2 Ideal measurements and the Lüders rule 74
- 9.3 Temporal correlations and certificates 75
- 9.4 Isoclinic subspaces and algebras of projections 75
- 9.5 Invariants of triangles and quadruples 76

9.6 Generation and speciality 76
9.7 Instruments and memory 77
9.8 Octonionic quantum mechanics 77
**10 Discussion** **79**
**A Octonion and Jordan conventions; normal forms** **81**
**B Proofs for Sections 3 and 4** **95**
**C Certificate details and conventions** **100**
**D The spinor realization** **111**
**E Proper subalgebras and proofs for Section 7** **114**
**F An algebraic proof of the obstruction** **119**
**G Formalization** **125**

# 1 Introduction

Exceptional algebraic structure can stay invisible in restricted descriptions and in short sequences of measurements, and become detectable once one description has to stay consistent under repeated ideal measurement. A precise form of this holds for the Albert algebra $J = H_3(\mathbb{O})$ of Hermitian $3 \times 3$ matrices over the octonions $\mathbb{O}$, with the symmetrised product $x \circ y = \frac{1}{2}(xy + yx)$. It is the one exceptional system that single-system reconstructions of quantum theory leave standing (Section 1.3).

Margin notes name the Lean declaration that proves the adjacent statement, with a remark where the formal route or scope differs; algebraic models are defined in Appendix G.

The statement has two halves. Take four ideal yes–no tests on $J$ that generate the algebra, meaning that sums and symmetrised products of the four tested elements and the unit $e$ reach all of $J$, and start from the maximally mixed state $e/3$, the analogue of the density matrix $I/3$. Through six tests in sequence, the Albert statistics are those of a random choice between two quaternionic quantum systems (Proposition 6.2). At seven they are not. A round trip, four tests run forward and then back to the first, has seven steps and measures a square: a random choice raises it by the variance between the two systems, while the Albert value is the square of their mean (paragraph "Why the mixture fails at seven" in Section 6.2). No Hilbert-space model with ideal measurements, of any dimension, matches every sequence of length seven (Theorem 6.1).

Behind both halves stands a criterion. A Hilbert-space model is asked only to reproduce probabilities, yet reproducing all of them forces a Jordan homomorphism, a linear map preserving the product $\circ$, from the algebra the tests generate into operators (Theorem 4.1).

Ideality, the requirement that each test disturb the state only as much as its answer requires (for sharp tests the Lüders rule, Theorem 3.3), is an assumption of the comparison and cannot be certified from the statistics. A quantum device whose tests disturb the state after answering reproduces every word probability of the Albert tests, however long the words (Theorem 8.1, Corollary 8.7), so no table of word probabilities shows that its tests were ideal. The results compare two model classes under the same postulate, ideal tests on the Albert algebra and ideal tests on Hilbert spaces, and Section 8 shows that the postulate does real work: without it the exceptional statistics have an ordinary quantum implementation. Nothing here bears on whether nature is quantum, and no complexity claim is made: $J$ is a 27-dimensional real vector space on which the ideal updates act as linear maps.

## 1.1 Four buttons and a light

Imagine a sealed box with four buttons and a light that answers yes or no. Each button is an ideal test of an *atom* of $J$. An atom plays the part that a rank-one projector plays in quantum theory: a yes–no question that cannot be split into finer ones, and equally the pure state in which its answer is certainly yes. The four atoms are chosen to generate $J$, and the box starts in the maximally mixed state $e/3$. Someone who has only the buttons can estimate the probability of every sequence of answers.

What "ideal" means is easiest to see on an ordinary quantum particle with three paths. Leave path three empty and let paths one and two interfere. Two

detectors can both answer "not on path three": one leaves the other two paths alone, the other quietly records which of them was taken. Only the first preserves the interference. An ideal test changes the state only as much as its answer requires, and in quantum theory it is the Lüders rule $\rho \mapsto P\rho P$ (Theorem 3.3).

An ordinary quantum device that may disturb the state after answering reproduces every sequence of the four Albert tests, however long, and its tests are even sharp and repeatable: each answer is given by a projection, and an immediate repetition confirms it (Corollary 8.7). Require each of its tests to make only the ideal change, and some prediction fails by the seventh press (Theorem 6.1).

The schedule matters too. Choose any seven-press schedule, a fixed sequence of seven of the four tests; an ordinary quantum device whose tests are ideal can be built to agree with the box on that schedule and on every run of six presses or fewer, and some other seven-press schedule must then give a different probability (Proposition 6.13(a) and Theorem 6.1).

## 1.2 A warm-up with a qubit and a qutrit

Take a qubit in the state $\rho$ and two yes–no measurements given by rank-one projections $P_1$, $P_2$ on $\mathbb{C}^2$, with transition probability $o = \mathrm{Tr}(P_1P_2)$. By the Lüders rule[1] the outcome "yes" leaves $P_1\rho P_1$ and has probability $\mathrm{Tr}(\rho P_1)$. Since $P_1P_2P_1 = o\,P_1$, a "yes" on $P_2$ followed by a "yes" on $P_1$ multiplies this probability by $o^2$:

$$\Pr(\mathrm{yes}_1, \mathrm{yes}_2, \mathrm{yes}_1) = \mathrm{Tr}(\rho P_1)\, o^2.$$

A "yes" on a rank-one projection forgets everything about the past except a number.

A "no" does not. On a qutrit with rank-one $P_1$, the outcome "no" updates by the rank-two projection $1 - P_1$ and leaves a two-dimensional piece of the state coherent. Three measurements on a qutrit, with "no" outcomes allowed, see the phase of the triangle formed by their unit vectors $a_1, a_2, a_3$. Writing $\Pr(i, \bar{\jmath}, k)$ for the probability of "yes" on $i$, "no" on $j$, "yes" on $k$, from the maximally mixed state, and $\Pr(i, k)$ for "yes" on $i$ then "yes" on $k$,

$$\mathrm{Re}\langle a_i|a_j\rangle\langle a_j|a_k\rangle\langle a_k|a_i\rangle = \tfrac{3}{2}\big[\Pr(i,k) + \Pr(i,j,k) - \Pr(i,\bar{\jmath},k)\big],$$

the real part of the Bargmann invariant of the three states[2]. The reason is that the anticommutator is a signed sum of the identity and the two Lüders updates, $\frac{1}{2}(PX + XP) = \frac{1}{2}\big(X + PXP - (1-P)X(1-P)\big)$.

In this sense sequences of ideal measurements measure Jordan products. A Jordan algebra is a real vector space with a commutative product $\circ$ of this kind (Section 3.1), and in it the identity reads $p \circ x = \frac{1}{2}(x + U_p x - U_{e-p}x)$, where $U_p$ and $U_{e-p}$ are the ideal updates (Section 3), and the right side makes sense for $J$.

[1] G. Lüders, Über die Zustandsänderung durch den Meßprozeß, *Ann. Phys. (Leipzig)* (6) **8** (1950) 322–328. doi:10.1002/andp.19504430510; English translation: Concerning the state-change due to the measurement process, *Ann. Phys. (Leipzig)* **15** (2006) 663–670, doi:10.1002/andp.20065180904.

[2] V. Bargmann, Note on Wigner's theorem on symmetry operations, *J. Math. Phys.* **5** (1964) 862–868. doi:10.1063/1.1704188.

To fit such statistics in a Hilbert space, fix finitely many measurements and ask for a Hilbert space, one projection $P_i$ per measurement and a state $\varphi$ such that every *word*, a finite sequence of outcomes in chronological order, has the same probability as in the system:

$$\varphi\big(Q_1 Q_2 \cdots Q_m \cdots Q_2 Q_1\big) = \Pr(q_1, \ldots, q_m), \qquad Q_t \in \{P_i, 1 - P_i\}.$$

These are the sequential projective representations of Budroni, Moroder, Kleinmann and Gühne[3], here allowed to be real, complex or quaternionic and to use any state (model class (I), Definition 3.4).

The rank of $P_i$ must be free: a comparison restricted to rank-one projections would declare the quaternionic qutrit, with its rank-two projections on $\mathbb{C}^6$, non-quantum.

Can ideal yes–no measurements on $J$ be modelled in this way, in any dimension and with projections of any rank? For four pure states that generate $J$, the answer depends on how the instruments are modelled:

| Instruments allowed in the quantum model | Reproduces all word probabilities? |
|---|---|
| One projection per test, Lüders update on both outcomes (class (I)) | no, in any dimension (Thm. 4.1) |
| One projection per test, arbitrary post-measurement channels | yes, on $\mathbb{C}^{138}$; on $\mathbb{C}^{136}$ from $e/3$ (Cor. 8.7) |
| Arbitrary completely positive instruments (class (II)) | yes, on $\mathbb{C}^{69}$; on $\mathbb{C}^{68}$ from $e/3$ (Thm. 8.1, Cor. 8.2) |

The exclusion is relative to class (I). A model with arbitrary instruments in infinite dimension reproduces any process (Section 3.4).

Asking less reveals more. If each yes–no test is replaced by a complete three-outcome measurement of a Jordan frame, three mutually orthogonal atoms summing to $e$ (the analogue of an orthonormal basis), the statistics become those of a classical Markov chain in the transition probabilities, and the exception disappears (Proposition 8.10).

### 1.3 The exception that single-system axioms leave standing

Reconstructions of quantum theory from postulates about a single system narrow the possibilities to the finite-dimensional formally real (also called Euclidean) Jordan algebras, in which a sum of squares vanishes only if every term does. Jordan, von Neumann and Wigner[4] classified them. The list consists of the real, complex and quaternionic quantum systems $H_n(\mathbb{R})$, $H_n(\mathbb{C})$, $H_n(\mathbb{H})$, the spin factors,

[3] C. Budroni, T. Moroder, M. Kleinmann and O. Gühne, Bounding temporal quantum correlations, *Phys. Rev. Lett.* **111** (2013) 020403. doi:10.1103/PhysRevLett.111.020403. arXiv:1302.6223.

[4] P. Jordan, J. von Neumann and E. Wigner, On an algebraic generalization of the quantum mechanical formalism, *Ann. of Math.* **35** (1934) 29–64. doi:10.2307/1968117.

whose state spaces are balls like the Bloch ball of a qubit, and one more: the 27-dimensional Albert algebra $J$, also called the octonionic qutrit[5]. It is *exceptional*: Albert[6] showed that it is not isomorphic to a Jordan algebra of self-adjoint operators on any Hilbert space. The classical algebraic witness of this is Glennie's identity[7], an equation of degree eight that holds in every special Jordan algebra, that is, in every Jordan algebra of operators, and fails in $J$. The known ways of removing $J$ use more than one system, dynamics, or the whole effect space, and the single-system interference tests that exist do not see it (Section 9.1).

The Albert algebra serves here as a *mathematical foil*, a consistent comparison case rather than a proposed physical theory. Every table, the list of all word probabilities, of ideal tests on a Euclidean Jordan algebra is a finite convex combination of finite-dimensional quantum tables and Albert tables (Remark 4.15); hence a finite linear inequality valid for ideal quantum measurements either holds for ideal tests on every Euclidean Jordan algebra or is violated by an Albert table. The Albert algebra is the unique simple exceptional ingredient in this decomposition, not a unique countermodel. Up to three tests cannot separate the classes, while four can (Corollary 4.13).

Glennie's identity separates $J$ from the special algebras by an equation. In the word statistics the first separation is an inequality. At four atoms the element $F$ obtained from Glennie's identity has squared length proportional to the invariant $\mathcal{A}$ of result 2 below, with a factor fixed by three of the atoms (Theorem F.6).

### 1.4 Results

A *test* is an ideal two-outcome measurement of an idempotent, an element $p$ with $p \circ p = p$ (the analogue of a projection). In a Euclidean Jordan algebra, a two-outcome instrument that leaves undisturbed every state on which its outcome is certain, and is sharp in the sense of Definition 3.2, is the Lüders instrument $(U_p, U_{e-p})$ of an idempotent; the same holds in quantum theory over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$ in any dimension (Theorem 3.3). The six transition probabilities of four atoms are $o_{ij} = \mathrm{tr}(p_i \circ p_j)$, with tr the trace of $J$ (Section 3.1).

**1.** *A sequential criterion for speciality.* Tests of atoms of $J$ have a sequential projective representation if and only if the atoms generate a proper subalgebra, and then from every state (Theorem 4.1). For general Euclidean Jordan algebras the criterion is that the state gives no weight to a summand isomorphic to $J$ in the algebra the tests generate, under a primitivity hypothesis on the tests (Theorem 4.11). An algebra is special if and only if every four of its tests are representable from every state (Corollary 4.13), and when the dimension is not

[5] J. C. Baez, The octonions, *Bull. Amer. Math. Soc.* **39** (2002) 145–205. doi:10.1090/S0273-0979-01-00934-X. arXiv:math/0105155; J. C. Baez and P. Schwahn, The Standard Model gauge group from the exceptional Jordan algebra, arXiv:2606.15235 (2026).

[6] A. A. Albert, On a certain algebra of quantum mechanics, *Ann. of Math.* (2) **35** (1934) 65–73. doi:10.2307/1968118.

[7] C. M. Glennie, Some identities valid in special Jordan algebras but not valid in all Jordan algebras, *Pacific J. Math.* **16** (1966) 47–59. doi:10.2140/pjm.1966.16.47.

fixed the statistics do not separate real, complex and quaternionic quantum theory (Proposition 4.14).

2. *Three tests fit a qutrit; four need not.* Any three atoms of $J$ lie in one complex subalgebra $W \cong H_3(\mathbb{C})$, so three tests are reproduced by an ordinary qutrit with rank-one projections (Proposition 5.1). For four atoms a single number $\mathcal{A} \geq 0$, built from thirteen traces of Jordan products, vanishes if and only if the atoms lie in a quaternionic subalgebra and is positive if and only if they generate $J$; in a normal form it is a squared octonion associator $(xy)z - x(yz)$ (Theorem 5.5). Seen from the subalgebra $W$ generated by three of the atoms, when it is complex (for instance when the triangle factor $\Delta_W$ of the three atoms is nonzero), $\mathcal{A} = \Delta_W \cdot N(P_W p_4)$, the triangle factor times the determinant $N$ of the fourth atom projected onto $W$ (Proposition 5.6). The maximum is $4/19683$, and at a maximiser each atom looks maximally mixed to the tests of the other three (Proposition 5.12, Corollary 5.14).

3. *The tetrahedron identity.* In every sequential projective representation of four tests, the relations forced by the statistics give Clifford relations among three triangle holonomies, operators such as $P_1P_2P_3P_1$ that go round a triangle of tests (the operator form of a Bargmann phase), and associativity of operator products then forces the tetrahedron statistic $\mathcal{D}$, a function of the probabilities of words of length at most four, to vanish (Theorem 5.9). For atoms of $J$, $\mathcal{D} \prod o_{ij} = \mathcal{A}$ (Proposition 5.10). A positive $\mathcal{D}$ alone is not a witness: for an explicit generating quadruple S (Section 6) a representation on $\mathbb{C}^{12}$ reproduces all thirteen traces (Proposition 6.16), and the companion paper gives a real representation with this property for every generating quadruple.

4. *Finite certificates.* For every four atoms that generate $J$, from the maximally mixed state $e/3$, words of length at most seven separate the octonionic word probabilities from every sequential projective representation (Theorem 6.1). Conversely a representation on $\mathbb{C}^{12}$, a mixture of two quaternionic ones, matches every word of length at most six, so the minimum separating length is $L_{\min} = 7$, and no linear equality on those words separates the table (Proposition 6.2). Every sequential projective representation misses some word probability of length at most seven by an amount bounded below by a positive constant that depends on the quadruple (Theorem 6.1(b)). For two explicit configurations M and S a single vector of forced negative squared norm suffices (Theorem 6.10). A companion paper, *Finite certificates for sequential measurements in the exceptional Jordan algebra*, gives explicit values of the constant for them, together with further finite-certificate details, robustness constants and bounds for five tests. An explicit list of 2,041 words of length at most eleven, the same for every generating quadruple, suffices from the maximally mixed state, and 8,164 words from every state (Theorem 6.15, whose proof is machine-checked). No linear equality separates through length five for any four atoms from any state either (Remark 6.18, proved in the companion paper and in Lean); the question at greater lengths is open.

5. *Five tests, every four quantum.* There are five atoms of $J$ such that every four of their tests have a sequential projective representation on $\mathbb{C}^6$, while the five together have none, in any dimension and from any state (Theorem 7.2). Four is the least number of tests that can show the obstruction, and five is the number that must be inspected in the worst case: a family of tests of atoms is representable as soon as every subfamily of at most five is (Proposition 7.4, Corollary 7.5); in the language of convexity, five is a Helly number. Words of length at most eleven exclude the five tests, by the length-seven argument for four atoms obtained from them by a reflection, carried back to the five tests; this bound is proved in the companion paper and machine-checked (Appendix G).

6. *Where the projection postulate enters.* With arbitrary completely positive instruments, the statistics of $n$ tests of atoms have an exact realization on $\mathbb{C}^{17n+1}$ (Theorem 8.1), and with sharp effects and arbitrary post-measurement channels one on $\mathbb{C}^{2(17n+1)}$ (Proposition 8.6, Corollary 8.7), a dimension that the companion paper lowers to $26n - 7$. So the hypothesis that does the work is the Lüders update on both outcomes, not projectivity of the effects. Conversely, with sharp effects, agreement on the words of length at most seven forces some test to disturb a state on which its answer is certain (Proposition 6.21), and instruments calibrated on their own yes and no subspaces to within a tolerance set by the margin of Theorem 6.1(b) (Section 6.6) miss the octonionic table on some word of length at most seven (Proposition 6.20 and (6.16)).

## 1.5 Relation to earlier work

The algebraic half of the criterion, that a JB-algebra (a complete normed generalisation of a Euclidean Jordan algebra) is special if and only if it admits no surjective Jordan homomorphism onto $J$, is due to Alfsen, Shultz and Størmer[8] and Roelands and Tiersma[9]; it is combined here with the operational half, in which matching statistics produce a Jordan homomorphism from the generated algebra into operators (Section 4.3).

That every triple of atoms lies in a complex subalgebra was asserted without proof by Ortega and Santander[10]; a proof is given in Section 5.1. The Clifford mechanism behind the tetrahedron identity is due to Wong and Wolf[11] for isoclinic subspaces; Section 5.4 adapts it to arbitrary representations.

[8] E. M. Alfsen, F. W. Shultz and E. Størmer, A Gelfand–Neumark theorem for Jordan algebras, *Adv. Math.* **28** (1978) 11–56. doi:10.1016/0001-8708(78)90044-0.

[9] M. Roelands and S. Tiersma, On the Shirshov–Cohn theorem for JB-algebras, arXiv:2605.24427v1 (2026). Thm. 2.11.

[10] R. Ortega and M. Santander, Trigonometry of 'complex Hermitian' type homogeneous symmetric spaces, *J. Phys. A* **35** (2002) 7877–7918. arXiv:math-ph/0112030.

[11] Y.-C. Wong, Isoclinic $n$-planes in Euclidean $2n$-space, Clifford parallels in elliptic $(2n-1)$-space, and the Hurwitz matrix equations, *Mem. Amer. Math. Soc.* **41** (1961). doi:10.1090/memo/0041; J. A. Wolf, Geodesic spheres in Grassmann manifolds, *Illinois J. Math.* **7** (1963) 425–446. doi:10.1215/ijm/1255644951; J. A. Wolf, Elliptic spaces in Grassmann manifolds, *Illinois J. Math.* **7** (1963) 447–462. doi:10.1215/ijm/1255644952. Correction: D. B. Shapiro and J. A. Wolf, On isoclinic closure: correction to "Elliptic spaces in Grassmann manifolds", *Illinois J. Math.* **19** (1975) 316–317. doi:10.1215/ijm/1256050819.

Albert's theorem[12], that $J$ is not a Jordan algebra of operators, enters only as the last step of the criterion (Theorem F.5). The step before it assumes only that a model with ideal tests on a Hilbert space reproduces the word probabilities from one state of $J$, in any dimension, with projections of any rank and any state of the model, and reconstructs from this a unital Jordan homomorphism from the algebra the tests generate into the operators on the cyclic (Gelfand–Naimark–Segal) space of the model's state (Proposition 4.8); for a general Euclidean Jordan algebra the exclusion holds only for states that give an Albert summand positive weight (Theorem 4.11). Albert's theorem says nothing about the length of words, about the Helly number five for families of tests (Corollary 7.5) or about the sharp-effect realization of Corollary 8.7, whose tests are not ideal. For Albert atoms the thirteen traces, and with them $\mathcal{A}$, in a normal form a squared octonion associator, are functions of the probabilities of words of length at most four, yet an ideal Hilbert-space model matches every word through length six from $e/3$ (Proposition 6.2), and seven is the first length at which the table excludes every such model (Theorem 6.1). The traces exclude a representation only through the purity relations, which involve longer words, and the model of Proposition 6.2 reproduces them as averages over two quaternionic blocks whose conditional states differ from the octonionic ones, a difference that first shows as a variance at length seven.

A good faith effort was made, using AI systems, to search for prior work; Section 9 describes the sources found and how each relates to the results here.

Section 2 sketches the proofs, and the external theorems used without proof are listed in Appendix A.1.

## 2 Sketch of the proofs

A Hilbert-space model is not asked to embed $J$, so the proof of Theorem 4.1 has to show that reproducing all word probabilities forces enough structure inside the model for Albert's theorem to apply. It has three steps, and the four-point case has a second, more explicit route.

**Matching statistics forces purity relations.** Let $q$ be an atom of $J$. After the outcome "yes" on $q$ the unnormalised state is a multiple of $q$, since the Lüders update satisfies $U_q x = \langle q, x\rangle\, q$, with the trace form $\langle q, x\rangle = \operatorname{tr}(q \circ x)$ (Section 3.1). For any later word $v = (v_1, \dots, v_m)$ put $c = \operatorname{tr}(U_{v_m} \cdots U_{v_1} q)$ and let $\check{v} = (v_{m-1}, \dots, v_1)$ be $v$ without its last letter, in reverse order. The fundamental formula $U_{U_a b} = U_a U_b U_a$ gives, for every history $t$ (a word performed before $q$),

$$\Pr(t, q, v, \check{v}, q) - 2c\, \Pr(t, q, v) + c^2 \Pr(t, q) = 0.$$

On the model side let $E_v$ be the sequential product of the projections of the outcomes of $v$, $P_t$ the product of the projections of the history, and $\Omega$ the cyclic vector of the cyclic representation of the state, the Hilbert space that the Gelfand–Naimark–Segal construction builds from the words and the state. In a sequential

[12] Albert, "On a certain algebra of quantum mechanics", note 6.

projective representation the same combination of the model's probabilities is the squared norm $\|z\, P_t\Omega\|^2$ with $z = P_q E_v P_q - c P_q$ (Lemmas 4.3 and 4.5). So matching three word probabilities kills one vector, and matching all words kills the operator $z$ (Lemma 4.4). The identities $z = 0$ are the purity relations.

**The relations rebuild a Jordan homomorphism.** The Peirce reflection $U_{2p_i - e}$, the counterpart of conjugation by $2P_i - 1$, is a Jordan automorphism and a signed sum of the identity and the two Lüders updates of test $i$; on the model side, conjugation by $2P_i - 1$ is the same signed sum (equation (4.4)). So the data control reflected projections as well as the original ones. Images of the atoms under words in the reflections span the algebra $B$ generated by the tests (Lemma 4.7). Two families of relations follow: purity relations $P_\alpha P_\beta P_\alpha = \lambda_{\alpha\beta} P_\alpha$ between reflected projections, and every linear relation among reflected atoms in $B$, transported to the projections. In the cyclic representation both families vanish, and the transport $\eta$ from $B$ to operators is a well-defined unital Jordan homomorphism with $\langle \Omega, \eta(x)\Omega\rangle = \langle \rho, x\rangle$ (Proposition 4.8).

**Exceptionality finishes.** If $B = J$, then $\eta$ is a Jordan homomorphism from $J$ into the associative algebra of bounded operators. Theorem F.5 (Appendix F), an algebraic form of Albert's theorem[13], gives $\eta = 0$, against $\eta(e) = 1$. Conversely, if the atoms generate a proper subalgebra, it is special and embeds in some $H_D(\mathbb{C})$. Inside the image the state has a unique trace representative, which is positive, so it is a density matrix that reproduces every word (Lemma 4.9); an order-unit Hahn–Banach extension gives a second proof (Appendix B.8). In a general Euclidean Jordan algebra the argument needs each test or its complement to be primitive in $B$, since that is what gives the purity relations (Theorem 4.11).

**The tetrahedron.** For four tests the purity relations with words of length at most three read $P_a P_b P_a = o_{ab} P_a$, $P_a(P_b P_c + P_c P_b)P_a = 2R_{abc} P_a$ and its three-letter analogue (R3) (Lemma 5.8), with the transition probabilities $o_{ab}$ and the traces $R_{abc} = \langle p_a \circ p_b, p_c\rangle$ of Definition 5.2. Fix a vertex $j$ of the tetrahedron of tests and form the three triangle holonomies, such as $h_1 = P_j P_k P_l P_j / \sqrt{o_{jk} o_{kl} o_{lj}}$, on the range of $P_j$.

The relations make each $h_a$ unitary with scalar Hermitian part $c_a$, the cosine of the triangle's Bargmann phase, and give the skew parts $N_a$ Clifford relations $N_a N_b + N_b N_a = -2M_{ab}$ for a matrix $M$ of traces. The product $h_1 h_2 h_3$ is, after two contractions, the holonomy of the far face, whose Hermitian part is fixed by a fourth cosine $c_4$. Expanding the product gives $c_4 = S + \Gamma$, with $S$ a scalar and $\Gamma$ the volume element of the Clifford relations, and regrouping products, which needs associativity, gives $\Gamma^2 = \det M$. Hence $\mathcal{D} = \det M - (c_4 - S)^2 = 0$ in every representation (Theorem 5.9).

[13] Ibid.

**The octonions break it.** In a normal form of four atoms of $J$, the three triangle phases become unit octonions $h_a = c_a + v_a$. Their imaginary parts satisfy the pairwise anticommutator relations $v_a v_b + v_b v_a = -2M_{ab}$, and the real part of $(h_1 h_2)h_3$ is the prescribed face scalar $c_4$; only the regrouping fails. By the Harvey–Lawson identity the Gram determinant of three imaginary octonions exceeds their squared triple product by a quarter of their squared associator, and $\mathcal{D} \prod o_{ij} = \mathcal{A}$ (Proposition 5.10, Theorem 5.5). So a positive $\mathcal{A}$ excludes every sequential projective representation of the statistics from the maximally mixed state $e/3$ (Corollary 5.11).

**Finitely many words.** Compactness turns the all-word exclusion into a finite linear inequality of unspecified length (Corollary 4.16). The relations are needed only on the vectors reached from $P_j\Omega$ by at most two holonomies. Recording every word that each use of a relation requires gives the list of 2,041 words of length at most eleven, which decides from the maximally mixed state (Theorem 6.15).

At length seven the relations hold on enough short histories to fix the Gram matrix of twenty-eight vectors such as $P_A P_z P_y P_x \xi$, where $\xi$ is the normalised state vector after a "yes" on $A$, as polynomials in the traces. The same Gram matrix arises in an explicit indefinite form built from the Clifford algebra of the holonomies with one reflection inserted, and there, using once the trace relation (the one polynomial relation among the thirteen traces, Proposition 5.3), three combinations have Gram determinant $-4\mathcal{A}^3/\big((o_{AB}o_{AC}o_{AD})^9 \prod_x (1 - o_{Ax})\big) < 0$. Thus length seven always suffices (Theorem 6.1).

# 3 Ideal tests and their sequences

On a Hilbert space the Lüders rule[14] replaces $\rho$ by $P\rho P$ on "yes" and by $(1-P)\rho(1-P)$ on "no" (unnormalised; the trace is the probability). It leaves unchanged every state on which the outcome is certain. Theorem 3.3 shows that this property, together with sharpness, forces the same rule in every Euclidean Jordan algebra.

## 3.1 Euclidean Jordan algebras in brief

A *Euclidean* (equivalently, formally real) *Jordan algebra* is a finite-dimensional real vector space $V$ with a commutative bilinear product $x \circ y$, unit $e$, the Jordan identity $(x \circ y) \circ (x \circ x) = x \circ (y \circ (x \circ x))$, and formal reality ($\sum_k x_k \circ x_k = 0$ only if all $x_k = 0$). An example is $H_n(\mathbb{C})$ with $x \circ y = \frac{1}{2}(xy + yx)$. A *Jordan homomorphism* between Jordan algebras is a linear map $\phi$ with $\phi(x \circ y) = \phi(x) \circ \phi(y)$, *unital* if it maps unit to unit; a bijective one from $V$ to itself is a *Jordan automorphism*. With the states and effects below it is a Jordan-algebraic system in the sense of[15]. The

[14] Lüders, "Über die Zustandsänderung durch den Meßprozeß", note 1.

[15] H. Barnum and A. Wilce, Local tomography and the Jordan structure of quantum theory, *Found. Phys.* **44** (2014) 192–212. doi:10.1007/s10701-014-9777-1. arXiv:1202.4513; H. Barnum, M. P. Müller and C. Ududec, Higher-order interference and single-system postulates characterizing quantum theory, *New J. Phys.* **16** (2014) 123029. doi:10.1088/1367-2630/16/12/123029. arXiv:1403.4147.

standard facts below are from Faraut and Korányi[16] and are listed with the other imported results in Appendix A.1.

*Spectra, trace, cone.* Every $x$ is $\sum_k \lambda_k c_k$ with distinct real $\lambda_k$ and pairwise orthogonal idempotents $c_k$ summing to $e$[17]. An idempotent is *primitive* if it is nonzero and not a sum of two nonzero orthogonal idempotents; a *Jordan frame* is a set of orthogonal primitive idempotents with sum $e$. Every $x$ is also $\sum_k \mu_k c_k$ over a Jordan frame $\{c_k\}$, the $\mu_k$ not necessarily distinct[18].

The trace $\operatorname{tr} x$ is the sum of the eigenvalues with multiplicity ($\operatorname{tr} q = 1$ for primitive $q$), and $\langle x, y\rangle = \operatorname{tr}(x \circ y)$ is an inner product with $\langle x \circ y, z\rangle = \langle y, x \circ z\rangle$. The cone $V_+$ of squares is closed, self-dual, has $e$ in its interior[19], and the trace is faithful on it.

*Effects* are the $f$ with $0 \le f \le e$, *states* the $\rho \ge 0$ with $\operatorname{tr}\rho = 1$, acting by $f \mapsto \langle \rho, f\rangle$; every effect counts as a possible outcome.

*Quadratic representation and Peirce spaces.* Put $U_a x = 2a \circ (a \circ x) - (a \circ a) \circ x$, so $U_a x = axa$ in $H_n(\mathbb{C})$. Each $U_a$ is self-adjoint and positive, and the fundamental formula $U_{U_a b} = U_a U_b U_a$ holds[20]. For an idempotent $p$, $V$ is the orthogonal sum of $V_\lambda(p) = \{x : p \circ x = \lambda x\}$, $\lambda \in \{1, \frac{1}{2}, 0\}$, with $V_1 \circ V_0 = 0$, $V_\lambda \circ V_{1/2} \subseteq V_{1/2}$ ($\lambda = 0, 1$), $V_{1/2} \circ V_{1/2} \subseteq V_1 + V_0$[21]. For $V = J$ these spaces are written $J_\lambda(p)$. Hence $U_{\alpha p + \beta(e-p)}$ acts on $V_1, V_{1/2}, V_0$ by $\alpha^2, \alpha\beta, \beta^2$: $U_p$, $U_{e-p}$ are the orthogonal projections onto $V_1(p)$, $V_0(p)$, and the Peirce reflection $U_{2p-e}$ (+1 on $V_1 \oplus V_0$, $-1$ on $V_{1/2}$) is a Jordan automorphism, the sign being a $\mathbb{Z}/2$-grading.

$V_1(p)$ is a Euclidean Jordan algebra with unit $p$, so a nonzero $p$ is primitive exactly when $V_1(p) = \mathbb{R}p$, and then

$$U_p x = \langle p, x\rangle\, p \qquad (p \text{ primitive},\ x \in V). \tag{3.1}$$

As a state update, $U_p$ is the *Lüders map* (a compression in the sense of Alfsen and Shultz[22], a filter in the sense of Wilce[23] and of Barnum, Müller and Ududec[24]).

*Classification and speciality.* Unital subalgebras are again Euclidean Jordan algebras; $\mathrm{Jord}\langle e, S\rangle$ is the one generated by $S$. Every Euclidean Jordan algebra is a direct sum of simple ideals, and the simple ones are $H_n(\mathbb{R})$, $H_n(\mathbb{C})$, $H_n(\mathbb{H})$, the spin factors $\mathbb{R} \oplus \mathbb{R}^m$ with product $(x_0, \mathbf{x}) \circ (y_0, \mathbf{y}) = (x_0 y_0 + \mathbf{x}\cdot\mathbf{y},\ x_0\mathbf{y} + y_0\mathbf{x})$, and the 27-dimensional Albert algebra $J = H_3(\mathbb{O})$ of Hermitian octonionic $3 \times 3$ matrices with $x \circ y = \frac{1}{2}(xy + yx)$ (Appendix A.3; Jordan, von Neumann and Wigner[25];[26]).

---

[16] J. Faraut and A. Korányi, *Analysis on Symmetric Cones*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 1994.

[17] Ibid. Thm. III.1.1.

[18] Ibid. Thm. III.1.2.

[19] Ibid. Thm. III.2.1.

[20] Ibid. Props. II.3.3, III.2.2.

[21] Ibid. Prop. IV.1.1, Thm. IV.2.1.

[22] E. M. Alfsen and F. W. Shultz, *Geometry of State Spaces of Operator Algebras*, Birkhäuser, Boston, 2003.

[23] A. Wilce, Conjugates, filters and quantum mechanics, *Quantum* **3** (2019) 158. doi:10.22331/q-2019-07-08-158. arXiv:1206.2897.

[24] Barnum, Müller and Ududec, *Higher-order interference and single-system postulates*, note 15.

[25] Jordan, von Neumann and Wigner, *On an algebraic generalization*, note 4.

[26] Faraut and Korányi, *Analysis on Symmetric Cones*, note 16. Prop. III.4.4, Thm. V.3.7.

A Jordan algebra is *special* if it is isomorphic to a Jordan subalgebra of $A^+$, an associative algebra $A$ with product $\frac{1}{2}(xy + yx)$ (for a Euclidean Jordan algebra, special means JC: isomorphic to a Jordan algebra of self-adjoint operators on a complex Hilbert space). All simple ones except $J$ are special; $J$ is not (Albert[27]).

On $J$, $\operatorname{tr} e = 3$, and $e/3$ is the *maximally mixed state*; the primitive idempotents are the *atoms*, or pure states, and form the octonionic projective plane $\mathbb{O}P^2$; $o_{ij} = \operatorname{tr}(p_i \circ p_j) \in [0, 1]$ is the transition probability between atoms $p_i$ and $p_j$.

*Formalized.* `AlbertChallenge.isAtom_iff_primitive`: trace-one idempotents are the primitive ones.

**Lemma 3.1** (Support)**.** *Let $c$ be an idempotent and $x \geq 0$ with $\langle x, c\rangle = 0$. Then $U_c x = 0$ and $x = U_{e-c}x$.*

*Narrower in Lean.* `native_positive_zero_pairing_support`. Stated for $V = J$ only.

*Proof.* $U_c x \geq 0$ and $\operatorname{tr} U_c x = \langle c, x\rangle = 0$, so $U_c x = 0$. Write $x = x_{1/2} + x_0$. For real $t \neq 0$, $U_{tc+(e-c)}x = t x_{1/2} + x_0 \geq 0$; dividing by $|t|$ and letting $t \to \pm\infty$ gives $\pm x_{1/2} \geq 0$, so $x_{1/2} = 0$. □

## 3.2 Ideal tests

A two-outcome *instrument* on $V$ is a pair $(T_1, T_0)$ of positive linear maps with $\operatorname{tr}(T_1\rho + T_0\rho) = \operatorname{tr}\rho$; outcome $k$ has probability $\operatorname{tr} T_k\rho$ and leaves $T_k\rho$. Its *measured effect* is $a = T_1^*(e)$, and $F_1 = \{\rho : \langle\rho, a\rangle = 1\}$, $F_0 = \{\rho : \langle\rho, a\rangle = 0\}$ (states) are the exposed faces on which outcome 1, resp. 0, is certain. Complete positivity is not assumed.

**Definition 3.2** (Ideal test)**.** A two-outcome instrument $(T_1, T_0)$ is an *ideal test* if (Id) (*ideality*) $T_k\rho = \rho$ for every $\rho \in F_k$, $k = 0, 1$; and (Sh) (*sharpness*) there are no state $\sigma$ and effect $f$ with $\langle\sigma, f\rangle = 1$ and $\langle\rho, f\rangle = 0$ for all $\rho \in F_0 \cup F_1$. The same definition applies to a real, complex or quaternionic Hilbert space $\mathcal{H}$ of any dimension, with positive real-linear maps on self-adjoint trace-class operators, density operators as states, effects $0 \leq f \leq 1$, and $a = T_1^*(1)$.

*Narrower in Lean.* `PositiveInstruments.IsIdealTest`. Defined for finite-dimensional positive forms; the trace-class Hilbert-space version is not defined.

Condition (Id) is ideality in the sense of Busch, Grabowski and Lahti[28], and Günaydin, Piron and Ruegg[29] call the octonionic case an ideal measurement of the first kind. As (Id) constrains the instrument only on the faces, (Sh) is needed (Example B.1). On a Euclidean Jordan algebra, (Sh) holds exactly when the measured effect is an idempotent (proof of Theorem 3.3(a)).

**Theorem 3.3** (Ideal tests are Lüders instruments)**.**

*Narrower in Lean.* `PositiveInstruments.idealTest_iff_quadraticRepresentation`. Part (a) only.

*(a) Let $V$ be a Euclidean Jordan algebra. A two-outcome instrument on $V$ is an ideal test if and only if $T_1 = U_p$ and $T_0 = U_{e-p}$ for an idempotent $p \in V$; then $p$ is the measured effect.*

*(b) Let $\mathcal{H}$ be a real, complex or quaternionic Hilbert space of any dimension. A two-outcome instrument on $\mathcal{H}$ is an ideal test if and only if $T_1X = PXP$ and $T_0X = (1-P)X(1-P)$ for an orthogonal projection $P$; then $P$ is the measured effect.*

[27] Albert, "On a certain algebra of quantum mechanics", note 6.

[28] P. Busch, M. Grabowski and P. J. Lahti, Repeatable measurements in quantum theory: their role and feasibility, *Found. Phys.* **25** (1995) 1239–1266. doi:10.1007/BF02055331.

[29] M. Günaydin, C. Piron and H. Ruegg, Moufang plane and octonionic quantum mechanics, *Commun. Math. Phys.* **61** (1978) 69–85. doi:10.1007/BF01609468.

Theorem 3.3 is a variant, with these operational hypotheses, of known characterisations: on a unital JB-algebra the pairs of bicomplementary positive projections are exactly the pairs $(U_p, U_{e-p})$ (Alfsen and Shultz[30], with hypotheses on kernels and ranges); on $B(\mathcal{H})$, ideality implies the Lüders rule under the hypotheses of Kleinmann[31], Chiribella and Yuan[32] and Cho[33]. These sources treat operator algebras; reconstructions of Jordan-algebraic systems take $U_p$ as an axiom[34].

*Proof of (a). The measured effect is an idempotent.* For $\rho \geq 0$, $\langle \rho, a\rangle = \operatorname{tr} T_1\rho \geq 0$ and $\langle \rho, e-a\rangle = \operatorname{tr} T_0\rho \geq 0$, so $0 \leq a \leq e$ by self-duality. Let $a = \sum_\lambda \lambda c_\lambda$ and $p = c_1$ ($p$ and $c_0$ possibly 0). For a state $\rho$, $\langle \rho, e-a\rangle = \sum_{\lambda<1}(1-\lambda)\langle \rho, c_\lambda\rangle$ is a sum of nonnegative terms, so $\rho \in F_1$ iff $\langle \rho, e-p\rangle = 0$, iff $\rho \in V_1(p)$ by Lemma 3.1; likewise $F_0$ is the set of states in $V_1(c_0)$. If $r = e - p - c_0 \neq 0$, the state $r/\operatorname{tr} r$ and the effect $r$ violate (Sh), since $r \in V_0(p) \cap V_0(c_0)$ is orthogonal to every state in $V_1(p)$ or $V_1(c_0)$. So $r = 0$ and $a = p$.

*$T_1$ on the Peirce spaces of $p$.* By (Id), $T_1$ fixes the states in $V_1(p)$, which span $V_1(p)$. For $y \in V_0(p)$, $y \geq 0$: $T_1 y \geq 0$ and $\operatorname{tr} T_1 y = \langle y, p\rangle = 0$, so $T_1 y = 0$; such $y$ span $V_0(p)$. For $z \in V_{1/2}(p)$ choose $\kappa > 0$ with $e + z/\kappa \geq 0$. For every real $t$,

$$0 \leq U_{p+t(e-p)}\big(e + z/\kappa\big) = p + \tfrac{t}{\kappa}\, z + t^2(e-p),$$

and applying $T_1$ gives $p + \frac{t}{\kappa} T_1 z \geq 0$; letting $t \to \pm\infty$, $T_1 z = 0$. So $T_1 = U_p$. The instrument $(T_0, T_1)$ satisfies the same hypotheses with measured effect $e - p$, so $T_0 = U_{e-p}$.

*Converse.* $U_p, U_{e-p}$ are positive, $U_p e + U_{e-p} e = e$, and they fix the states of $V_1(p)$ and $V_0(p)$, which are $F_1$ and $F_0$; so (Id) holds. An effect vanishing on these states has $\langle p, f\rangle = \langle e-p, f\rangle = 0$, so $\operatorname{tr} f = 0$ and $f = 0$; so (Sh) holds. □

Part (b) is the same argument with spectral projections (Appendix B.1). It gives the Lüders, not the von Neumann, rule for degenerate $P$, and repeatability, that an immediate repetition of the test confirms its outcome, follows from $U_p^2 = U_p$.

### 3.3 Words and their probabilities

Fix $V$, idempotents $p_1, \ldots, p_n$ (the *tests*; a test is identified with its idempotent) and a state $\rho$. Test $i$ has outcome *yes*$_i$ with update $U_{p_i}$ and *no*$_i$ with update $U_{e-p_i}$. A *word* (an outcome sequence) is a finite sequence $w = (q_1, \ldots, q_m)$, $q_t \in \{p_i, e-p_i\}$, in chronological order, with probability

$$\Pr_\rho(w) = \operatorname{tr}\big(U_{q_m} \cdots U_{q_1}\rho\big) = \langle \rho, \theta(w)\rangle, \qquad \theta(w) = U_{q_1} U_{q_2} \cdots U_{q_m} e. \tag{3.2}$$

[30] Alfsen and Shultz, *Geometry of State Spaces*, note 22. Thms. 2.83, 7.12.

[31] M. Kleinmann, Sequences of projective measurements in generalized probabilistic models, *J. Phys. A* **47** (2014) 455304. doi:10.1088/1751-8113/47/45/455304. arXiv:1402.3583. Sec. III A.

[32] G. Chiribella and X. Yuan, Bridging the gap between general probabilistic theories and the device-independent framework for nonlocality and contextuality, *Inf. Comput.* **250** (2016) 15–49. doi:10.1016/j.ic.2016.02.006. arXiv:1504.02395.

[33] K. Cho, *Effectuses in Categorical Quantum Foundations*, PhD thesis, Radboud University Nijmegen, 2019. arXiv:1910.12198. Thm. 6.3.34.

[34] Barnum, Müller and Ududec, *Higher-order interference and single-system postulates*, note 15.

The effect $\theta(w)$ is the iterated sequential product of the outcomes in the sense of Gudder and Greechie[35] (on a Euclidean Jordan algebra an invariant continuous sequential product is $U_{\sqrt{a}}b$[36]); the empty word has $\theta = e$. A skipped test is absent from the word. Summing over the outcomes of the last test gives the shorter word, but summing over a test in the middle does not, since $U_p + U_{e-p}$ removes the $V_{1/2}(p)$ component. A *table* is a finite labelled family of tests together with all its word probabilities.

Every $\theta(w)$ lies in $B = \mathrm{Jord}\langle e, p_1, \ldots, p_n\rangle$, so these sequential statistics (also called temporal correlations[37]) depend only on $B$, the tests, and $\rho|_B$. If all letters are atoms, (3.1) gives $\Pr_\rho(w) = \langle \rho, q_1\rangle \prod_t \langle q_t, q_{t+1}\rangle$.

On a Hilbert space, projections $P_i$ and the Lüders rule give

$$\Pr(w) = \mathrm{Tr}(Q_m \cdots Q_1 \varrho Q_1 \cdots Q_m) = \mathrm{Tr}(\varrho E_w),$$

with

$$E_w = Q_1 \cdots Q_{m-1} Q_m Q_{m-1} \cdots Q_1, \qquad Q_t = P_i\,(\mathrm{yes}_i), \quad Q_t = 1 - P_i\,(\mathrm{no}_i),$$

the counterpart of $\theta(w)$, since $U_Q X = QXQ$.

**Other measurements.** The test of $e - p$ is that of $p$ with outcomes relabelled. A *complete measurement* is the three-outcome Lüders instrument $(U_{f_1}, U_{f_2}, U_{f_3})$ of a Jordan frame of $J$; reading only whether the outcome was $k$ updates by $\sum_{l\neq k} U_{f_l} \neq U_{e-f_k}$, so it is not a family of tests (Proposition 8.10). A *generalised Lüders instrument* has non-idempotent effects $a_k$ and the square-root update $x \mapsto U_{a_k^{1/2}} x$[38].

## 3.4 Three model classes

Every theorem names one of the following three model classes; class (I) is the comparison class of every main theorem.

**Definition 3.4** (Sequential projective representation; class (I))**.** Let $V$, tests $p_1, \ldots, p_n$ and a state $\rho$ be given. A *sequential projective representation* of their statistics consists of a Hilbert space $\mathcal{H}$ over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$ of arbitrary (possibly infinite, possibly nonseparable) dimension, one orthogonal projection $P_i$ on $\mathcal{H}$ for each test, of arbitrary rank, and a state $\varphi$ of the representation, that is, a positive real-linear functional on the bounded self-adjoint operators with $\varphi(1) = 1$ (not necessarily normal, faithful or pure), such that

$$\varphi(E_w) = \Pr_\rho(w) \qquad \text{for every word } w. \tag{3.3}$$

*Formalized.* `HilbertProjectiveData` ($\mathbb{R}$, $\mathbb{C}$), `QuaternionicProjectiveData`; the proofs use the wider class `AlgebraicProjectiveData`.

[35] S. Gudder and R. Greechie, Sequential products on effect algebras, *Rep. Math. Phys.* **49** (2002) 87–111. doi:10.1016/S0034-4877(02)80007-6.

[36] J. van de Wetering, Three characterisations of the sequential product, *J. Math. Phys.* **59** (2018) 082202. doi:10.1063/1.5031089. arXiv:1803.08453.

[37] Budroni et al., "Bounding temporal quantum correlations", note 3; G. Vitagliano and C. Budroni, Leggett–Garg macrorealism and temporal correlations, *Phys. Rev. A* **107** (2023) 040101. doi:10.1103/PhysRevA.107.040101. arXiv:2212.11616.

[38] P. Busch and J. Singh, Lüders theorem for unsharp quantum measurements, *Phys. Lett. A* **249** (1998) 10–12. doi:10.1016/S0375-9601(98)00704-X. arXiv:1304.0054.

If (3.3) is required only for a given set of words, it is a sequential projective representation *of those words*.

The same projection is used each time a test is performed, and both outcomes update by the Lüders rule. For complex spaces and density-operator states these are the sequential projective quantum representations of Budroni, Moroder, Kleinmann and Gühne[39], who showed that membership on a finite set of words is decided, in arbitrary dimension, by a moment-matrix condition. Real and quaternionic spaces and non-normal states give the same data (Lemma 4.3 and Appendix B.10), so their name is used here. A degeneracy-breaking rule is not in the class; with such rules dimension-dependent bounds appear[40].

*Remark* 3.5 (A fixed unitary between tests). Let $\mathcal{H}$ be finite-dimensional, with projections $P_i$, a density operator $\varrho$ and a unitary $Y$ applied before each test, so that a word has probability $\mathrm{Tr}(Q_m Y \cdots Q_1 Y \varrho Y^* Q_1 \cdots Y^* Q_m)$. If these probabilities equal $\Pr_\rho(w)$ for every word, the table has a sequential projective representation without $Y$. For a word $u$ let $\sigma_u$ be the unnormalised operator that meets the next test after $u$ ($\sigma_\varnothing = Y\varrho Y^*$), and $K$ the span of the ranges of all $\sigma_u$. In every Euclidean Jordan algebra $U_{e-p}U_p = 0$, so a "yes" and then a "no" on the same test, or the reverse, has probability zero after every $u$. In the model the first is $\|(1-P_i)YP_i\sigma_u^{1/2}\|_2^2$ (Hilbert–Schmidt norm), so $(1-P_i)YP_i$ and likewise $P_iY(1-P_i)$ vanish on $K$, that is $P_iY = YP_i$ on $K$. The range of $\sigma_{(u,q)}$ is $YQ$ applied to that of $\sigma_u$, so $YP_iK \subseteq K$ and $Y(1-P_i)K \subseteq K$; hence $YK \subseteq K$, so $YK = K$ in finite dimension, and $P_iK = Y^{-1}YP_iK \subseteq K$. On $K$ the unitary commutes with every $Q_t$, so $Q_mY\cdots YQ_1 = Y^{m-1}Q_m\cdots Q_1$ there, and invariance of the trace gives the table of $(K, P_i|_K, \sigma_\varnothing)$, a sequential projective representation. The step $YK = K$ uses finite dimension; infinite dimension is not treated here.

**Definition 3.6** (Completely positive realization; class (II))**.** Let a finite family of instruments be given, with one letter per outcome. A *completely positive realization* of dimension $d$, in the sense of Monràs and Winter[41] (also called a finite quantum memory), is a density matrix $\sigma$ on $\mathbb{C}^d$ and one completely positive map $\mathcal{E}_a$ on $M_d(\mathbb{C})$ per letter, the same each time the letter occurs, such that the outcome maps of each instrument sum to a trace-preserving map and $\mathrm{Tr}(\mathcal{E}_{a_m}\cdots\mathcal{E}_{a_1}\sigma) = \Pr(a_1\cdots a_m)$ for every word. For tests the letters are $p_i$, $e-p_i$, the right side is $\Pr_\rho(w)$, and a skipped test is absent from the word.

*Formalized.* `IsNormalizedFinite CPModel`, `RealFiniteKrausModel`.

A finite-dimensional complex sequential projective representation is a completely positive realization ($X \mapsto P_iXP_i$, $X \mapsto (1-P_i)X(1-P_i)$), and a real or quaternionic one becomes one after passing to $\mathbb{C}$ (Proposition 4.14(a)). An infinite-dimensional one is not, and a completely positive realization need not use projections or Lüders updates. So excluding one class does not exclude the

[39] Budroni et al., "Bounding temporal quantum correlations", note 3.

[40] C. Budroni and C. Emary, Temporal quantum correlations and Leggett–Garg inequalities in multilevel systems, *Phys. Rev. Lett.* **113** (2014) 050401. doi:10.1103/PhysRevLett.113.050401. arXiv:1309.3678.

[41] A. Monràs and A. Winter, Quantum learning of classical stochastic processes: the completely-positive realization problem, *J. Math. Phys.* **57** (2016) 015219. doi:10.1063/1.4936935. arXiv:1412.3634.

other: the four-atom statistics excluded from class (I) have a completely positive realization on $\mathbb{C}^{69}$ (Theorem 8.1).

*Class (III): arbitrary instruments in infinite dimension.* A model always exists: a classical register stores the current normalised state of $V$ and emits each outcome with its conditional probability. Statements about class (III) are constructions, never exclusions.

**Terminology.** Terms defined later: *B-primitive* tests (Definition 4.10); the *purity operator* and *purity relations* (Lemma 4.5); *orbit labels* (Section 4.3); the *four-point invariant* $\mathcal{A}$ ((5.4)) and the *double-pair traces* (Definition 5.2); the *tetrahedron identity* and the *tetrahedron statistic* $\mathcal{D}$ (Theorem 5.9); the *triangle holonomies* (Section 5.4); the *minimum separating word length* (Section 6); *projective equalities* (Section 6.5); and *line subalgebras* (Section 7.3).

# 4 A sequential criterion for speciality

When do the word probabilities of Section 3 have a sequential projective representation (Definition 3.4)? For atoms of the Albert algebra the answer is: exactly when the atoms generate a proper subalgebra of $J$.

## 4.1 Statement for the Albert algebra

**Theorem 4.1** (The Albert algebra)**.** *Let $p_1, \dots, p_n$ be atoms of $J = H_3(\mathbb{O})$ and $B = \mathrm{Jord}\langle e, p_1, \dots, p_n\rangle$. The model class is (I). The following are equivalent.*

*(1) $B$ is special; equivalently, $B \neq J$.*

*(2) For every state $\rho$ of $J$ there is a sequential projective representation on $\mathbb{C}^d$ for some finite $d$, with a density-matrix state.*

*(3) For at least one state $\rho$ of $J$ there is a sequential projective representation, over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$, of any dimension, with projections of any rank and any state $\varphi$.*

*If $B = J$ there is no sequential projective representation from any state. Moreover there are then a length $L$ and a linear inequality in the probabilities of words of length at most $L$ that holds in every sequential projective representation and is violated by the statistics of $p_1, \dots, p_n$ from every state of $J$.*

*Different form in Lean.* `AlbertChallenge.albert_criterion`: (1) as an embedding of a subalgebra containing the atoms into complex Hermitian matrices; (2) with $d \in \{6, 24\}$ and projections fixed for all states; (3) on spaces in Lean's lowest universe, algebraic models admitted. (3) in any universe, and the last two sentences: `AlbertChallenge.no_representation_of_generates`, `AlbertChallenge.linear_separator_of_generates`.

The same holds if some tests are complements $e - p$ of atoms, since the test of $e - p$ is the test of $p$ with outcomes relabelled and $B$ does not change. Four atoms can generate $J$ (Theorem 5.5 gives an exact criterion), while three never do (Proposition 5.1).

*Remark* 4.2 (Nothing non-associative happens in time). The update maps $U_{p_i}$ and $U_{e-p_i}$ are linear maps of the 27-dimensional real vector space $J$. They compose associatively, and the probability (3.2) of a word is obtained by composing them in the order of the outcomes, so no "non-associativity in time" is observed. The non-associative Jordan product enters only through *which* linear maps the updates are (for a model without the Lüders requirement, see Theorem 8.1).

## 4.2 The mechanism: null vectors on every history

Fix a sequential projective representation $(\mathcal{H}, P_i, \varphi)$. The first step keeps only what the statistics can see.

**Lemma 4.3** (Real cyclic representation)**.** *Let $A_0$ be the real $*$-algebra spanned by the finite products of $P_1, \dots, P_n$ (the empty product 1 included). Put $\varphi_s(X) = \varphi\big(\frac{1}{2}(X + X^*)\big)$ and $\langle X, Y\rangle_\varphi = \varphi_s(X^*Y)$. This form is symmetric and positive semidefinite; let $\mathcal{H}_\varphi$ be the real Hilbert space obtained by quotienting its null space and completing, and $\Omega = [1]$. Left multiplication by $X \in A_0$ descends to a bounded operator $\pi(X)$ on $\mathcal{H}_\varphi$, $\pi$ is a real $*$-representation, and each $\pi(P_i)$ is an orthogonal projection. For a word $w = (q_1, \dots, q_m)$ put $P_w = Q_m \cdots Q_1$ (last outcome leftmost). Then*

$$\Pr(w) = \varphi(E_w) = \big\|\pi(P_w)\Omega\big\|^2,$$

*and the vectors $\pi(P_w)\Omega$, $w$ ranging over all words, span a dense subspace of $\mathcal{H}_\varphi$.*

*Formalized.* `cyclicHilbertModel`, `cyclicHilbertModel_probability`, starting from an algebraic model.

*Proof.* This is the Gelfand–Naimark–Segal construction, written so that $\varphi$ is evaluated only on self-adjoint operators; it therefore covers real and quaternionic spaces and non-normal or non-faithful states. Since each $Q_t$ is a self-adjoint idempotent, $E_w = P_w^* P_w$, so $\varphi(E_w) = \varphi_s(P_w^* P_w) = \|\pi(P_w)\Omega\|^2$. Every finite product of the $P_i$ is $P_w$ for an all-"yes" word $w$, which gives density. Boundedness and the $*$-property are routine (Appendix B.2). □

The data enter through squares. A *formal word combination* is a finite real combination $X = \sum_w c_w[w]$ of word symbols. It has two evaluations: the operator $E(X) = \sum_w c_w E_w \in A_0$ in the representation, and the effect $\theta(X) = \sum_w c_w \theta(w) \in B$ in $V$. Matching all words means exactly $\varphi(E(X)) = \langle \rho, \theta(X)\rangle$ for every $X$. Prefixing a word $t = (t_1, \dots, t_k)$ to every word of $X$ gives $t \cdot X$, and

$$E(t \cdot X) = P_t^*\, E(X)\, P_t, \qquad \theta(t \cdot X) = U_{t_1} \cdots U_{t_k}\, \theta(X). \tag{4.1}$$

**Lemma 4.4** (Vanishing moments give null vectors on every history)**.** *Let $Z \in A_0$ be self-adjoint and $X$ a formal word combination with $E(X) = Z^2$ as an identity of operators and $\theta(X) = 0$. Let $\mathcal{T}$ be a set of words (histories). If the representation reproduces $\Pr(t\,w)$ for every $t \in \mathcal{T}$ and every $w$ in the support of $X$, then $\pi(Z)\pi(P_t)\Omega = 0$ for every $t \in \mathcal{T}$. If it reproduces all words, then $\pi(Z) = 0$.*

*Narrower in Lean.* `allHistoryNullContract`: the null-set part only, left invariance of the null vectors and extension from word vectors to their span; the step from $E(X) = Z^2$, $\theta(X) = 0$ and matching to $\pi(Z)\pi(P_t)\Omega = 0$ is not part of this declaration.

*Proof.* By Lemma 4.3, (4.1) and matching,

$$\begin{aligned}\|\pi(Z)\pi(P_t)\Omega\|^2 &= \varphi(P_t^* Z^2 P_t) = \varphi(E(t \cdot X)) \\ &= \sum_w c_w \Pr(t\,w) = \langle \rho, U_{t_1} \cdots U_{t_k} \theta(X)\rangle = 0.\end{aligned}$$

If all words are reproduced, the bounded operator $\pi(Z)$ vanishes on a dense subspace. □

The hypothesis puts $ZP_t$ in the null left ideal of the moment functional, in the language of Pironio, Navascués and Acín[42] (Section 9.3). The quantifier over histories matters: when $\varphi$ is not faithful, one vanishing expectation says nothing about the operator, and finite arguments (Section 6) must record which histories they use.

Vanishing effects $\theta(X) = 0$ come from purity: a "yes" on an atom leaves a multiple of that atom, and in the algebra this is the identity

$$U_q x = \lambda_q(x)\, q, \qquad \lambda_q(x) = \langle q, x\rangle / \langle q, q\rangle \qquad (x \in B), \tag{4.2}$$

valid for every idempotent $q$ that is primitive in $B$, because then $U_q B = B_1(q) = \mathbb{R}q$. For an atom $q$ of $J$, $\langle q, q\rangle = 1$ and (4.2) is (3.1).

**Lemma 4.5** (Purity)**.** *Let $q$ be one of the outcomes $p_i$ or $e - p_i$, primitive in $B$ (for instance an atom of $J$), and $P_q$ the corresponding projection $P_i$ or $1 - P_i$. Let $v = (v_1, \ldots, v_m)$ be a word, $c = \lambda_q(\theta(v))$ (for an atom, $c = \mathrm{tr}(U_{v_m} \cdots U_{v_1} q)$), and*

$$z_q(v) = P_q\, E_v\, P_q - c\, P_q.$$

*For every word $t$,*

$$\left\| \pi(z_q(v))\pi(P_t)\Omega \right\|^2 = \Pr(t, q, v, \check{v}, q) - 2c\, \Pr(t, q, v) + c^2 \Pr(t, q), \tag{4.3}$$

*with the probabilities of the representation on the right and $\check{v} = (v_{m-1}, \ldots, v_1)$; the same combination of probabilities in $V$ is zero. Hence if the representation reproduces these three words, $\pi(z_q(v))\pi(P_t)\Omega = 0$; if it reproduces all words, $\pi(z_q(v)) = 0$.*

*Narrower in Lean.* `purityResidual_threeHistories`; the vanishing in $V$, for outcomes that are atoms of $J$: `NativeAnalysis.native_corner_long_probability`.

The operator $z_q(v)$ is called the purity operator and $\pi(z_q(v)) = 0$ the purity relations.

*Proof.* Expanding, $z_q(v)^2 = P_q E_v P_q E_v P_q - 2c P_q E_v P_q + c^2 P_q$, and these three operators are the sequential products $E$ of the words $(q, v, \check{v}, q)$, $(q, v)$ and $(q)$; so $z_q(v)^2 = E(X)$ for $X = [q, v, \check{v}, q] - 2c[q, v] + c^2[q]$, and (4.3) is Lemma 4.4's computation.

*The effect vanishes.* In $V$, iterating the fundamental formula gives

$$U_{v_1} \cdots U_{v_{m-1}} U_{v_m} U_{v_{m-1}} \cdots U_{v_1} = U_{\theta(v)},$$

hence, by (4.2),

$$\theta(q, v, \check{v}, q) = U_q U_{\theta(v)} U_q e = U_{U_q \theta(v)} e = U_{cq} e = c^2 q, \qquad \theta(q, v) = U_q \theta(v) = c\, q,$$

so $\theta(X) = c^2 q - 2c^2 q + c^2 q = 0$. For an atom, $c = \langle q, \theta(v)\rangle = \mathrm{tr}(U_{v_m} \cdots U_{v_1} q)$ because the $U$'s are self-adjoint. No conditional probability is divided by a possibly zero number. □

Lemma 4.5 feeds the certificates of Section 6; the reconstruction needs it only for $v$ of length one, transported along reflections.

[42] S. Pironio, M. Navascués and A. Acín, Convergent relaxations of polynomial optimization problems with non-commuting variables, *SIAM J. Optim.* **20** (2010) 2157–2180. doi:10.1137/090760155. arXiv:0903.4368.

### 4.3 Reconstruction of a Jordan homomorphism

Assume now that a sequential projective representation reproduces all words, and that every test $p_i$ is primitive in $B$; for atoms of $J$ with $B = J$ this holds. The construction builds a unital Jordan homomorphism from $B$ into the bounded self-adjoint operators $B(\mathcal{H}_\varphi)_{\mathrm{sa}}$ out of operators the data control, using no property of the tests beyond (4.2).

*Reflections are controlled by the data.* Put $r_i = 2p_i - e$ and $\Sigma_i = 2P_i - 1$. The Peirce reflection $U_{r_i}$ is a Jordan automorphism of $B$ (Section 3.1), and it is a combination of the two Lüders updates; so is conjugation by $\Sigma_i$:

$$U_{r_i} = 2(U_{p_i} + U_{e-p_i}) - \mathrm{id}, \qquad \Sigma_i Y \Sigma_i = 2P_i Y P_i + 2(1-P_i)Y(1-P_i) - Y. \tag{4.4}$$

*Orbit labels.* An *orbit label* is a pair $\alpha = (g, i)$ with $g = (g_1, \dots, g_k)$ a finite sequence of indices; it names

$$q_\alpha = U_{r_{g_1}} \cdots U_{r_{g_k}}\, p_i \in B, \qquad P_\alpha = \Sigma_{g_1} \cdots \Sigma_{g_k}\, P_i\, \Sigma_{g_k} \cdots \Sigma_{g_1} \in A_0.$$

Each *orbit point* $q_\alpha$ is primitive in $B$, as the image of a primitive idempotent under an automorphism, and $P_\alpha$ is a projection. Write $r_\alpha = 2q_\alpha - e$ and $\Sigma_\alpha = 2P_\alpha - 1$. Different labels may name the same $q_\alpha$ but, a priori, different $P_\alpha$, until the linear relations below identify them. Composing the conjugations by $\Sigma_j$ gives (Appendix B.4):

**Lemma 4.6** (Lifts of conjugated products)**.** *For orbit labels $\alpha_1, \dots, \alpha_s$ there is a formal word combination $X$, a* lift*, with*

$$E(X) = P_{\alpha_1} \cdots P_{\alpha_{s-1}} P_{\alpha_s} P_{\alpha_{s-1}} \cdots P_{\alpha_1}, \qquad \theta(X) = U_{q_{\alpha_1}} \cdots U_{q_{\alpha_{s-1}}}\, q_{\alpha_s}.$$

*Moreover, if $E(X) = Y$ and $\theta(X) = y$, there is $X'$ with $E(X') = \Sigma_\alpha Y \Sigma_\alpha$ and $\theta(X') = U_{r_\alpha} y$. These identities hold for any projections $P_i$.*

*Formalized.* `hasSequentialEffectLift_orbit_corner`, `hasSequentialEffectLift_reflection`.

*The orbit spans $B$.* The Peirce decomposition gives, for an idempotent $q$ and any $x$, and in any associative algebra for a projection $P$ and any $Y$,

$$q \circ x = U_q x + \tfrac{1}{4}\big(x - U_{r_q} x\big) \quad (r_q = 2q - e), \tag{4.5}$$

$$P \circ Y = PYP + \tfrac{1}{4}\big(Y - \Sigma Y \Sigma\big) \quad (\Sigma = 2P - 1), \tag{4.6}$$

where $P \circ Y = \frac{1}{2}(PY + YP)$. (Both sides of (4.5) agree on each Peirce space of $q$.)

**Lemma 4.7** (Orbit span)**.** $B = \mathrm{span}\big(\{e\} \cup \{q_\alpha\}\big)$.

*Narrower in Lean.* `nativeOrbitSpan_iff_generated`, for tests that are atoms of $J$.

*Proof.* The span $B'$ of the right side lies in $B$ and is mapped into itself by each $U_{q_\alpha}$ (by (4.2)) and each $U_{r_\alpha}$ (which fixes $e$ and permutes orbit points), hence by each $q_\alpha \circ$ (by (4.5)). So $B'$ is a unital subalgebra containing every $p_i$, and $B' = B$. Details are in Appendix B.5. □

*Two families of null operators.* Let $\lambda_{\alpha\beta} = \lambda_{q_\alpha}(q_\beta)$, so $U_{q_\alpha} q_\beta = \lambda_{\alpha\beta} q_\alpha$.

- *Purity relations* (transported). $T_{\alpha\beta} = P_\alpha P_\beta P_\alpha - \lambda_{\alpha\beta} P_\alpha$. By Lemma 4.6, $T_{\alpha\beta}^2$ has a lift with effect $U_{q_\alpha} U_{q_\beta} q_\alpha - 2\lambda_{\alpha\beta} U_{q_\alpha} q_\beta + \lambda_{\alpha\beta}^2 q_\alpha$ (Appendix B.4), and this is 0 because the fundamental formula gives $U_{q_\alpha} U_{q_\beta} q_\alpha = U_{q_\alpha} U_{q_\beta} U_{q_\alpha} e = U_{\lambda_{\alpha\beta} q_\alpha} e = \lambda_{\alpha\beta}^2 q_\alpha$. For $\alpha = (\varnothing, i)$, $\beta = (\varnothing, j)$ this is Lemma 4.5 with $q = p_i$, $v = (p_j)$.
- *Linear relations* (transported). Whenever $c_0 e + \sum_r c_r q_{\alpha_r} = 0$ in $B$, put $Z = c_0 1 + \sum_r c_r P_{\alpha_r}$. Writing each $P_{\alpha_r} \circ P_{\alpha_s}$ in $Z^2$ by (4.6) and applying Lemma 4.6 gives a lift of $Z^2$ whose effect is, by (4.5), $\big(c_0 e + \sum_r c_r q_{\alpha_r}\big)^2 = 0$ (Appendix B.4).

By Lemma 4.4, $\pi(T_{\alpha\beta}) = 0$ and $\pi(Z) = 0$ for all such operators. In particular (a linear relation with $c_0 = 0$) $\pi(P_\alpha) = \pi(P_\beta)$ whenever $q_\alpha = q_\beta$.

**Proposition 4.8** (The reconstructed homomorphism)**.** *Suppose a sequential projective representation reproduces all words and every test is primitive in $B$. Then*

$$\eta\Big(c_0 e + \sum_r c_r q_{\alpha_r}\Big) = c_0\, 1 + \sum_r c_r\, \pi(P_{\alpha_r})$$

*defines a unital Jordan homomorphism $\eta : B \to B(\mathcal{H}_\varphi)_{\mathrm{sa}}$ with $\langle \Omega, \eta(x)\Omega\rangle = \langle \rho, x\rangle$ for all $x \in B$.*

*Narrower in Lean.* `exists_generatedCyclicJordanMap`. Stated for tests that are atoms of $J$, with $\eta$ valued in operators on the cyclic space of an algebraic model.

*Proof. Well defined.* By Lemma 4.7 every $x \in B$ has such an expression, and two expressions differ by a linear relation, which the transported linear relations send to 0: so $\eta$ is well defined, linear and unital.

*Reflections and corners.* It intertwines reflections, $\eta(U_{r_\alpha} x) = \pi(\Sigma_\alpha)\eta(x)\pi(\Sigma_\alpha)$, because $U_{r_\alpha}$ acts on labels as conjugation by $\Sigma_\alpha$ acts on the $P$'s, and the purity relations give the corner relation $\pi(P_\alpha)\eta(x)\pi(P_\alpha) = \lambda_{q_\alpha}(x)\pi(P_\alpha)$ (both checked in Appendix B.6).

*Products.* Now (4.5), (4.2) and (4.6) give the decisive line:

$$\eta(q_\alpha \circ x) = \lambda_{q_\alpha}(x)\,\pi(P_\alpha) + \tfrac{1}{4}\big(\eta(x) - \pi(\Sigma_\alpha)\eta(x)\pi(\Sigma_\alpha)\big) = \pi(P_\alpha) \circ \eta(x).$$

Both sides of $\eta(y \circ x) = \eta(y) \circ \eta(x)$ are linear in $y$, and $B$ is spanned by $e$ and the $q_\alpha$; so $\eta$ is a Jordan homomorphism. Finally $\langle \Omega, \pi(P_\alpha)\Omega\rangle = \varphi(E(X)) = \langle \rho, q_\alpha\rangle$ for the lift $X$ of $P_\alpha$ (Lemma 4.6), by matching. □

*The contradiction.* Let $B = J$. The map $\eta$ is a Jordan homomorphism into the real associative algebra of bounded operators on $\mathcal{H}_\varphi$. Theorem F.5 gives $\eta = 0$, whereas $\eta(e) = 1 \neq 0$ because $\|\Omega\| = 1$; equivalently, since $J$ is simple, $\eta$ would be a faithful representation of $J$ and its Jordan product by operators. This contradiction uses the algebraic proof in Appendix F; Albert's original theorem remains the historical source of exceptionality[43].

The argument used no four-point invariant, no finite rank, no normality or faithfulness of $\varphi$, no positivity of $\rho$, and nothing about the original Hilbert space beyond its cyclic representation.

[43] Albert, "On a certain algebra of quantum mechanics", note 6.

### 4.4 The converse: discard, embed, construct

Let $\rho$ be a state of $V$ and suppose $B = \mathrm{Jord}\langle e, p_1, \dots, p_n\rangle$ has no summand isomorphic to $J$ given positive weight by $\rho$ (Section 4.5 makes this precise). For atoms of $J$ with $B \neq J$ there is nothing to discard: $B$ has dimension below 27, so by the classification all its simple summands are special.

*Discard.* If $A \cong J$ is a summand of $B$ with unit $c$ and $\langle \rho, c\rangle = 0$, then $|\langle \rho, x\rangle| \leq \|x\| \langle \rho, c\rangle = 0$ for $x \in A$ (Appendix B.8). So $\mathrm{Pr}_\rho(w) = \langle \rho, \theta_{\mathrm{sp}}(w)\rangle$, where $\theta_{\mathrm{sp}}(w)$ is the image of $\theta(w)$ under the central projection $\pi_{\mathrm{sp}}$ of $B$ onto the sum $B_{\mathrm{sp}}$ of its special summands; $\pi_{\mathrm{sp}}$ is a unital Jordan homomorphism, so $\theta_{\mathrm{sp}}(w)$ is the sequential product of the tests $\pi_{\mathrm{sp}}(p_i)$.

*Embed.* Each special simple summand has a faithful unital representation on some $\mathbb{C}^{d_j}$, and their direct sum is a faithful unital Jordan homomorphism $\iota : B_{\mathrm{sp}} \to H_d(\mathbb{C})$ with $x \geq 0$ iff $\iota(x) \geq 0$ (Appendix B.8).

*Construct the state.* The following lemma gives the density matrix directly.

**Lemma 4.9** (Explicit state)**.** *Let $W$ be a Euclidean Jordan algebra with unit $e_W$, $\omega$ a linear functional on $W$ with $\omega(x^2) \geq 0$ for all $x \in W$ and $\omega(e_W) = 1$, $p_1, \dots, p_n$ idempotents of $W$, and $\iota : W \to H_d(\mathbb{C})$ an injective unital Jordan homomorphism. There is a unique $z \in W$ with*

$$\mathrm{Tr}\big(\iota(z)\iota(x)\big) = \omega(x) \qquad \text{for all } x \in W.$$

*Then $\sigma = \iota(z)$ is a density matrix, and the projections $P_i = \iota(p_i)$ on $\mathbb{C}^d$ with the state $X \mapsto \mathrm{Tr}(\sigma X)$ satisfy $\mathrm{Tr}(\sigma E_w) = \omega(\theta_W(w))$ for every word $w$, where $\theta_W(w) = U_{q_1} \cdots U_{q_m} e_W$ with $q_t \in \{p_i, e_W - p_i\}$. In particular, if $W$ is a unital subalgebra of a Euclidean Jordan algebra $V$ and $\omega(x) = \langle \rho, x\rangle$ for a state $\rho$ of $V$, they form a sequential projective representation of the statistics of the tests from $\rho$: $\mathrm{Tr}(\sigma E_w) = \mathrm{Pr}_\rho(w)$.*

*Narrower in Lean.* `FaithfulHermitianJordanEmbedding.exists_unique_positive_trace_representative`, `suppliedEmbedding` `StateTransport`. The density matrix is obtained, as printed, from a positive functional with value 1 at the unit; the word transport is stated only for unital subalgebras of $J$ and states of $J$.

*Proof. The element $z$.* The form $\tau(x, y) = \mathrm{Tr}(\iota(x)\iota(y))$ on $W$ is symmetric and bilinear, and positive definite: for $x \neq 0$, $\iota(x)$ is a nonzero Hermitian matrix and $\mathrm{Tr}(\iota(x)^2) > 0$. So the linear functional $\omega$ on the finite-dimensional space $W$ is $\tau(z, \cdot)$ for a unique $z$.

*$\sigma$ is a density matrix.* Write $z = \sum_k \mu_k c_k$ with distinct real $\mu_k$ and pairwise orthogonal nonzero idempotents $c_k$ of $W$ summing to $e_W$ (the spectral theorem in the Euclidean Jordan algebra $W$, Section 3.1). Each $\iota(c_k)$ is a Hermitian idempotent, an orthogonal projection, and is nonzero, so $\mathrm{Tr}\,\iota(c_k) \geq 1$. For $k \neq l$, $\iota(c_k)\iota(c_l) + \iota(c_l)\iota(c_k) = 2\iota(c_k \circ c_l) = 0$; for orthogonal projections $P, Q$ with $PQ + QP = 0$, multiplying by $P$ on the left and on the right gives $PQ + PQP = 0 = PQP + QP$, so $PQ = QP$ and then $PQ = 0$. Hence

$$\omega(c_k) = \tau(z, c_k) = \sum_l \mu_l \,\mathrm{Tr}\big(\iota(c_l)\iota(c_k)\big) = \mu_k \,\mathrm{Tr}\,\iota(c_k).$$

The left side is $\omega(c_k^2) \geq 0$. So every $\mu_k \geq 0$ and $\sigma = \sum_k \mu_k \iota(c_k) \geq 0$. Since $\iota$ is unital, $\mathrm{Tr}\,\sigma = \tau(z, e_W) = \omega(e_W) = 1$.

*The words.* In an associative algebra $2X \circ (X \circ Y) - X^2 \circ Y = XYX$, so a Jordan homomorphism satisfies $\iota(U_a b) = \iota(a)\iota(b)\iota(a)$. Unitality gives $\iota(e_W - p_i) = 1 - P_i$, so $\iota(q_t) = Q_t$ for every letter. Every $\theta_W(w)$ lies in $W$, and

$$\iota(\theta_W(w)) = \iota(U_{q_1} \cdots U_{q_m} e_W) = Q_1 \cdots Q_m \, 1 \, Q_m \cdots Q_1 = E_w,$$

using $Q_m^2 = Q_m$. Therefore $\mathrm{Tr}(\sigma E_w) = \tau(z, \theta_W(w)) = \omega(\theta_W(w))$. If $W$ is a unital subalgebra of $V$ and $\omega = \langle \rho, \cdot \rangle$, then $\theta_W(w) = \theta(w)$ and $\omega(\theta(w)) = \Pr_\rho(w)$. □

Apply the lemma to $W = B_{\mathrm{sp}}$, with unit $c_{\mathrm{sp}} = \pi_{\mathrm{sp}}(e)$, the idempotents $\pi_{\mathrm{sp}}(p_i)$, the embedding $\iota$ above and $\omega(x) = \langle \rho, x \rangle$ for $x \in B_{\mathrm{sp}}$. Then $\omega(x^2) \geq 0$, and $\omega(c_{\mathrm{sp}}) = \langle \rho, e \rangle = 1$ because the discarded summands have weight zero. Since $\pi_{\mathrm{sp}}$ is a unital Jordan homomorphism, $\theta_{B_{\mathrm{sp}}}(w) = \theta_{\mathrm{sp}}(w)$. With $P_i = \iota(\pi_{\mathrm{sp}}(p_i))$ and the density matrix of the lemma, $\mathrm{Tr}(\sigma E_w) = \langle \rho, \theta_{\mathrm{sp}}(w) \rangle = \Pr_\rho(w)$. Some $P_i$ may be 0 or 1.

*Proof of Theorem 4.1.* In (1), $B \neq J$ implies that $B$ is special by the classification, and $B = J$ is not special by Theorem F.5. (1) $\Rightarrow$ (2): embed and construct the state. (2) $\Rightarrow$ (3) is trivial. (3) $\Rightarrow$ (1): if $B = J$, the atoms are primitive in $B$, and a representation reproducing all words gives the unital homomorphism $\eta$ of Proposition 4.8, contradicting Theorem F.5. The finite inequality is Corollary 4.16(a). □

## 4.5 General Euclidean Jordan algebras

Let $V$ be a Euclidean Jordan algebra, $p_1, \ldots, p_n$ idempotents, $\rho$ a state, and $B = \mathrm{Jord}\langle e, p_1, \ldots, p_n \rangle$. As a Euclidean Jordan algebra,

$$B = B_{\mathrm{sp}} \oplus A_1 \oplus \cdots \oplus A_m, \tag{4.7}$$

where $B_{\mathrm{sp}}$ is the sum of the simple summands not isomorphic to $J$ and each $A_k \cong J$ (an *Albert summand*), with unit $c_k$, a central idempotent of $B$. The state $\rho$ *gives positive weight* to $A_k$ if $\langle \rho, c_k \rangle > 0$, and zero weight otherwise.

**Definition 4.10** ($B$-primitive test)**.** A test $p \notin \{0, e\}$ is *$B$-primitive* if $p$ or $e - p$ is a primitive idempotent of $B$.

This holds whenever $p$ or $e - p$ is primitive in $V$, so every test of a primitive idempotent of $V$ is $B$-primitive. It holds for every nontrivial idempotent when $V$ has rank at most three (e.g. $H_3(\mathbb{K})$, $J$, spin factors). It can fail when $V$ has rank at least four (Remark 4.12).

**Theorem 4.11** (Sequential criterion for speciality)**.** *Let $V$ be a Euclidean Jordan algebra, $p_1, \ldots, p_n \in V$ idempotents with generated algebra $B$ as in* (4.7)*, and $\rho$ a state of $V$. The model class is (I).* *Not formalized.*

*(a) If $\rho$ gives zero weight to every Albert summand of $B$ (in particular, if $B$ is special), there is a sequential projective representation on $\mathbb{C}^d$ for some finite $d$, with a density-matrix state.*

*(b) Suppose every test $p_i \notin \{0, e\}$ is $B$-primitive. If $\rho$ gives positive weight to some Albert summand of $B$, there is no sequential projective representation: no Hilbert space over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$, of any dimension, with projections of any rank and any state $\varphi$, reproduces all word probabilities.*

*Thus, for $B$-primitive tests, a sequential projective representation exists if and only if $\langle \rho, c_k \rangle = 0$ for every Albert summand $A_k$ of $B$.*

Part (a) is the construction of Section 4.4. Part (b) is the proof of Theorem 4.1 with three changes, carried out in Appendix B.7: tests $p_i \in \{0, e\}$ are dropped and $p_i$ is replaced by $e - p_i$ where necessary, so that every test is primitive in $B$; Lemma 4.5 and the reconstruction then use only (4.2); and injectivity is proved on one summand $A_k$ of positive weight, using $\langle \Omega, \eta(c_k)\Omega \rangle = \langle \rho, c_k \rangle > 0$. Unlike Theorem 4.1, part (b) depends on the state: when $B$ has an Albert summand and another summand, the same tests can have a representation from one state and none from another.

*Remark* 4.12 (Mixtures of exceptional systems). It remains unresolved here whether $B$-primitivity can be dropped from Theorem 4.11(b). The simplest open case: in $V = J \oplus J$ take tests $p_i = (a_i, a_i')$ with $\{a_i\}$ and $\{a_i'\}$ atoms each generating $J$ and the pairs generating $V$, and $\rho = \lambda\rho_1 \oplus (1-\lambda)\rho_2$ with $0 < \lambda < 1$. Each $p_i$ has rank two in $B = V$ and is not $B$-primitive, so (4.2) and the purity relations fail. The table is a mixture of two tables that each have no representation (Theorem 4.1), but the set of representable tables is convex, so this alone decides nothing.

## 4.6 Consequences

**Corollary 4.13** (Four tests decide; two or three never detect)**.** *For a Euclidean Jordan algebra $V$ and class (I), the following are equivalent: (i) $V$ is special; (ii) every finite family of tests of $V$ has a sequential projective representation from every state; (iii) every four tests of $V$ have a sequential projective representation from every state. In every $V$, any two and any three tests have a sequential projective representation on some $\mathbb{C}^d$ from every state.*

*Not formalized.*

*Proof.* If $V$ is special, every generated algebra is special, so Theorem 4.11(a) gives (i)⇒(ii) (Appendix B.9). If $V$ is not special it has a summand $A \cong J$, with unit $c$. Four atoms generating $A$ (Theorem 5.5) generate $B = A \oplus \mathbb{R}(e - c)$, in which they are primitive, and the state $c/3$ gives $A$ positive weight, so Theorem 4.11(b) applies. Three tests always generate a special algebra, by Proposition 5.1 (see also Section 9.6), so Theorem 4.11(a) applies (Appendix B.9). □

A family need not be representable when each of its four-element subfamilies is (Theorem 7.2). Let $Q(\mathbb{K})$, respectively $Q_{\mathrm{fin}}(\mathbb{K})$, be the set of tables produced by sequential projective representations over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$ in arbitrary, respectively finite, dimension, and $\mathcal{J}(V)$ the set of tables produced by tests of $V$ from all its states.

**Proposition 4.14** (Real, complex and quaternionic representations give the same tables)**.**

*(a)* $\mathcal{Q}_{\text{fin}}(\mathbb{R}) = \mathcal{Q}_{\text{fin}}(\mathbb{C}) = \mathcal{Q}_{\text{fin}}(\mathbb{H})$ *and* $\mathcal{Q}(\mathbb{R}) = \mathcal{Q}(\mathbb{C}) = \mathcal{Q}(\mathbb{H})$.

*(b) For every spin factor* $V$, $\mathcal{J}(V) \subseteq \mathcal{Q}_{\text{fin}}(\mathbb{C})$, *and the union of these sets over all spin factors is strictly smaller than* $\mathcal{Q}_{\text{fin}}(\mathbb{C})$.

*(c)* $\mathcal{J}(V) \subseteq \mathcal{Q}(\mathbb{C})$ *iff* $\mathcal{J}(V) \subseteq \mathcal{Q}_{\text{fin}}(\mathbb{C})$ *iff* $V$ *is special.*

*Proof.* Part (a) follows from two standard constructions: the cyclic representation of Lemma 4.3 turns every representation into a real one, finite-dimensional if the original is (Appendix B.2), and extension of scalars from $\mathbb{R}$ gives the converse. The rest follows from Theorem 4.11(a), Corollary 4.13 and one explicit table (Appendix B.10). □

*Narrower in Lean.* (a) every inclusion, by constructions preserving all word probabilities and finite dimension: `HilbertProjectiveData.complexExtensionModel_probability`, `HilbertProjectiveData.quaternionExtensionModel_probability`, `cyclicHilbertModel_probability`, `HilbertProjectiveData.cyclicHilbert_finite`, `QuaternionicProjectiveData.cyclicHilbert_finite_of_real`, `complexification_finiteDimensional`, `quaternionification_finiteDimensional`; the equalities are not stated as such. (b) and (c) are not formalized.

*Remark* 4.15 (The Albert algebra as a mathematical foil). Fix the number $n$ of tests. Write $\mathcal{Q}_n^{\text{fin}}$ for the finite-dimensional quantum tables above, $\mathcal{J}_n(V)$ for the tables of $n$ ideal tests on $V$ from all its states, and $\mathcal{J}_n = \bigcup_V \mathcal{J}_n(V)$, with $V$ ranging over finite-dimensional Euclidean Jordan algebras. Then

$$\mathcal{J}_n = \operatorname{conv}\big(\mathcal{Q}_n^{\text{fin}} \cup \mathcal{J}_n(J)\big), \tag{4.8}$$

where conv means finite convex combinations, without closure. Thus a linear inequality involving finitely many word probabilities, including the empty word if a constant term is needed, that holds on $\mathcal{Q}_n^{\text{fin}}$ holds on all of $\mathcal{J}_n$ if and only if it holds on $\mathcal{J}_n(J)$.

Indeed, decompose $V = \bigoplus_k V_k$ into simple ideals with units $c_k$, and let $\omega(x) = \langle \rho, x \rangle$ be the initial state functional. Put $\lambda_k = \omega(c_k)$. These weights are nonnegative and sum to one. For $\lambda_k > 0$ the functional $\omega_k = \omega|_{V_k}/\lambda_k$ is positive and takes value one at $c_k$, hence is a state of $V_k$ in its own trace normalization. Every quadratic update acts componentwise, so the word table is $\sum_k \lambda_k f_k$, where $f_k$ is the table of the component idempotents from $\omega_k$; zero-weight terms vanish. By the classification (Section 3.1), each $V_k$ is special or isomorphic to $J$. In the special case Theorem 4.11(a) gives $f_k \in \mathcal{Q}_n^{\text{fin}}$. Conversely, $\mathcal{Q}_n^{\text{fin}} \subseteq \mathcal{J}_n$, since complex Hermitian matrices are Euclidean Jordan algebras, and direct sums of algebras, tests and weighted states realize every finite convex combination. This proves (4.8); linearity gives the inequality assertion.

The unique *simple* exceptional algebra supplies the only additional simple ingredient. Several Albert tables may be needed in the convex combination; neither their replacement by one Albert table nor a unique countermodel is asserted. The decomposition does not decide whether a particular mixture of Albert tables is quantum (Remark 4.12).

**Corollary 4.16** (Finite state-independent inequalities)**.**

*Narrower in Lean.* `HilbertModelSeparation.uniformHilbertSeparator_generating_atoms`. Part (a) only.

*(a) In Theorem 4.1 with $B = J$ there are $L$, a linear functional $\Lambda$ on the probabilities of words of length at most $L$ (the empty word included) and $\gamma > 0$ such that $\Lambda \geq 0$ on every sequential projective representation and $\Lambda \leq -\gamma$ on the statistics from every state of $J$.*

*(b) In Theorem 4.11(b), let $c_{\mathrm{ex}} = \sum_k c_k$. For each $\varepsilon > 0$ the same holds for all states with $\langle \rho, c_{\mathrm{ex}} \rangle \geq \varepsilon$, with $L, \Lambda, \gamma$ allowed to depend on $\varepsilon$. If $B$ has an Albert summand and $c_{\mathrm{ex}} \neq e$, no single choice of $L$, $\Lambda$ and $\gamma > 0$ serves all states with $\langle \rho, c_{\mathrm{ex}} \rangle > 0$.*

*Proof.* The moment functionals of all representations form a compact convex set (Lemma 4.3, Tychonoff's theorem). If every length admitted a match, a nested intersection would give an exact all-word match, which Theorems 4.1 and 4.11 exclude. Strict separation of disjoint compact convex sets in $\mathbb{R}^N$ then gives $\Lambda$ (Appendix B.11). □

Compactness gives no value of $L$, $\gamma$ or the coefficients; Section 6 gives the lengths seven and eleven for four atoms.

# 5 Three atoms fit a qutrit; four need not

By Theorem 4.1, the algebra generated by the tests decides whether a sequential projective representation reproduces them. One atom $p$ generates $\mathbb{R}p \oplus \mathbb{R}(e-p)$, and a qubit reproduces its test. Two atoms $p, q$ generate a subalgebra of $\mathrm{span}\{e, p, q, p \circ q\}$, a copy of $\mathbb{R} \oplus H_2(\mathbb{R})$ when they are distinct and not orthogonal, and a real qutrit with two rank-one projections at the right angle reproduces them.

Conventions for octonions and for $J$ are those of Appendix A, and $F_4 = \mathrm{Aut}(J)$ is the automorphism group of $J$, a compact Lie group of dimension 52 (Appendix A.5). A *complex subalgebra* is a unital Jordan subalgebra of $J$ isomorphic to $H_3(\mathbb{C})$, and a *quaternionic subalgebra* one isomorphic to $H_3(\mathbb{H})$. The standard ones are $H_3(\mathbb{C}_i)$ and $H_3(\mathbb{H})$, with $\mathbb{C}_i = \mathrm{span}\{1, i\}$ and $\mathbb{H} = \mathrm{span}\{1, i, j, ij\}$.

## 5.1 Three atoms always sit in a qutrit

**Proposition 5.1** (Three atoms)**.** *Any three atoms $p_1, p_2, p_3$ of $J$ lie in a common complex subalgebra $W \subseteq J$. Consequently, for every state $\rho$ of $J$, the statistics of the three tests from $\rho$ have a sequential projective representation (class (I)) on $\mathbb{C}^3$ with rank-one projections and a density-matrix state.*

*Different route.* `complexTripleContract_holds`: the normalizing automorphisms are constructed in the development, without the facts on $F_4$ used in the printed proof. With only the space $\mathbb{C}^3$ recorded: `AlbertChallenge.three_atoms`.

Earlier results on triangles and on algebras generated by three idempotents are recalled in Sections 9.5 and 9.6. The proof below also covers degenerate triples, and it gives rank-one projections, where a quaternionic subalgebra would give only rank-two projections on $\mathbb{C}^6$.

*Proof.* Lemma A.3 gives $g \in F_4$ with $gp_1, gp_2, gp_3 \in H_3(\mathbb{C}_i)$; put $W = g^{-1}H_3(\mathbb{C}_i)$. The consequence is Lemma A.2: a unital Jordan isomorphism $\psi : W \to H_3(\mathbb{C})$

preserves the trace form, so it sends atoms to rank-one projections, and for the trace-orthogonal projection $\rho_W$ of $\rho$ onto $W$, $\psi(\rho_W)$ is a density matrix that reproduces every word in the three tests. □

## 5.2 Thirteen traces

Four atoms modulo $F_4$ have $4 \cdot 16 - 52 = 12$ parameters: $\mathbb{O}P^2$ has dimension 16, and when the four atoms generate $J$ (which is the generic case, Theorem 5.5) an automorphism fixing them fixes every Jordan product of them and so is the identity. Traces of Jordan products serve as coordinates.

**Definition 5.2** (The thirteen traces)**.** For atoms $p_A, p_B, p_C, p_D$ of $J$ define the six transition probabilities, the four shape invariants and the three double-pair traces

*Formalized.* `fpOverlap`, `fpTriangle`, `fpDoublePair1`, `fpOmega1` and their analogues.

$$o_{ij} = \langle p_i, p_j \rangle, \qquad R_{ijk} = \langle p_i \circ p_j, p_k \rangle,$$
$$t_1 = \langle p_A \circ p_B, p_C \circ p_D \rangle, \quad t_2 = \langle p_A \circ p_C, p_B \circ p_D \rangle, \quad t_3 = \langle p_A \circ p_D, p_B \circ p_C \rangle,$$

and the combinations $\omega_r = t_1 + t_2 + t_3 - 2t_r$ $(r = 1, 2, 3)$.

The shape invariants are symmetric in their labels, because the trace form is associative, $\langle x \circ y, z \rangle = \langle x, y \circ z \rangle$. For atoms of a qutrit, $p_k = |a_k\rangle\langle a_k|$, one has $o_{ij} = |\langle a_i | a_j \rangle|^2$ and $R_{ijk} = \mathrm{Re}\langle a_i | a_j \rangle\langle a_j | a_k \rangle\langle a_k | a_i \rangle$, the real part of the Bargmann invariant, which Brehm calls the shape invariant of the triangle[44]; so $R_{ijk}/\sqrt{o_{ij}o_{jk}o_{ki}}$ is the cosine of the Bargmann (geometric) phase. By Proposition 5.1 the same holds for any three atoms of $J$, computed in a complex subalgebra containing them. The double-pair traces carry the four-point data.

The trace of a Jordan triple product against an atom is also useful: by associativity of the trace form,

$$\begin{aligned} Q_{j;klm} &:= \langle p_j, (p_k \circ p_l) \circ p_m + (p_m \circ p_l) \circ p_k - (p_k \circ p_m) \circ p_l \rangle \\ &= \langle p_j \circ p_m, p_k \circ p_l \rangle + \langle p_j \circ p_k, p_m \circ p_l \rangle - \langle p_j \circ p_l, p_k \circ p_m \rangle, \end{aligned} \tag{5.1}$$

so that $Q_{A;BCD} = \omega_2$, $Q_{A;CDB} = \omega_3$ and $Q_{A;DBC} = \omega_1$.

**Reading the traces from words.** Start in the maximally mixed state $e/3$, write words in chronological order with $i$ for "yes" on test $i$ and $\bar{\imath}$ for "no". The Jordan multiplication operator $L_p x = p \circ x$ of an idempotent is a signed sum of the identity and the two Lüders updates,

$$L_p = \tfrac{1}{2}\big(\mathrm{id} + U_p - U_{e-p}\big), \tag{5.2}$$

since both sides act as $1, \frac{1}{2}, 0$ on the Peirce spaces of $p$; the same identity holds for a projection $P$ acting on operators by $X \mapsto \frac{1}{2}(PX + XP)$. So the operators generated by the Lüders updates are those generated by the $L_{p_i}$, and right-normed products $p_{i_1} \circ (p_{i_2} \circ (\cdots))$ paired with a state are combinations of word probabilities.

[44] U. Brehm, The shape invariant of triangles and trigonometry in two-point homogeneous spaces, *Geom. Dedicata* **33** (1990) 59–76. doi:10.1007/BF00147601.

Since $U_p e = p$ and $U_p x = \langle p, x\rangle\, p$ for an atom, a word that starts with $i$ and ends with $k$ has probability $\frac{1}{3}\langle p_i, X p_k\rangle$, where $X$ is the product of the intermediate $U$-operators, and (5.2) gives

$$o_{ij} = 3\Pr(i,j), \qquad R_{ijk} = \tfrac{3}{2}\big[\Pr(i,k) + \Pr(i,j,k) - \Pr(i,\bar{\jmath},k)\big],$$
$$t_1 = \tfrac{3}{4}\textstyle\sum_{\sigma,\tau} \varepsilon_\sigma \varepsilon_\tau \Pr(A, \sigma_B, \tau_C, D),$$

where $\sigma, \tau$ range over "skip", "yes", "no" with signs $+, +, -$; $t_2, t_3$ follow by relabelling. Transition probabilities need words of length two, shape invariants length three, double-pair traces length four.

The thirteen traces are not independent. Because $J$ has rank three, the fourth elementary symmetric function of any spectrum vanishes, and this gives one relation, linear in the shape invariants and double-pair traces.

**Proposition 5.3** (The trace relation)**.** *Let $V$ be a Euclidean Jordan algebra of rank three in which primitive idempotents have trace one (for instance $J$, or a complex or quaternionic subalgebra of $J$). For any four primitive idempotents $p_A, p_B, p_C, p_D$ of $V$,*

$$t_1+t_2+t_3 = \tfrac{1}{2} - \tfrac{1}{2}\sum_{i<j} o_{ij} + \tfrac{1}{2}\big(o_{AB}o_{CD} + o_{AC}o_{BD} + o_{AD}o_{BC}\big) + R_{ABC} + R_{ABD} + R_{ACD} + R_{BCD}. \tag{5.3}$$

*Narrower in Lean.* `rankThree_fourIdempotent_trace_relation`, `albert_fourAtom_trace_relation`. Stated for an explicit rank-three cubic model and for $J$; general rank-three Euclidean Jordan algebras are not treated.

*Proof.* For $x = \sum_i \lambda_i p_i$ with real $\lambda_i$, the spectral theorem[45] and power associativity make $s_k = \operatorname{tr}(x^k)$ the power sums of three eigenvalues, so Newton's identities give the polynomial identity in $\lambda$

$$24 e_4 = s_1^4 - 6 s_1^2 s_2 + 3 s_2^2 + 8 s_1 s_3 - 6 s_4 = 0.$$

The coefficient of $\lambda_A\lambda_B\lambda_C\lambda_D$ in this identity is (5.3); Appendix A.7 computes it. □

The relation is the four-point analogue of $N(p_1+p_2+p_3) = 1 - o_{12} - o_{13} - o_{23} + 2R_{123}$, the multilinear coefficient of the cubic norm $N$, the determinant of $J$ (Appendix A.3). It also follows by multiplying the generic minimal equation $x^3 - s_1 x^2 + e_2 x - e_3 e = 0$ by $x$ and taking the trace.

*Remark* 5.4 (Exactly one relation, and what kind). (a) Every polynomial relation among the thirteen traces of four atoms of $J$ is a multiple of (5.3), and the traces carry twelve parameters, matching the count above. A triangular local inverse realizes an open set of twelve trace coordinates; polynomial division by (5.3) then proves the assertion (Appendix A.7). No claim is made that they generate the $F_4$-invariants of four atoms.

(b) The relation needs rank three; it is not a consequence of the Jordan identity. For four rank-one projections of $H_n(\mathbb{C})$, $n \geq 4$, the same computation keeps $e_4$, and the left side minus the right side of (5.3) is $-\frac{1}{2}$ times the Gram determinant of the four unit vectors.

(c) It is not a witness. It is quadratic in word probabilities, and a general sequential projective representation need not satisfy it: four orthogonal rank-one projections on $\mathbb{C}^4$ with the maximally mixed state give $0 = \frac{1}{2}$. A representation that reproduces the words of length at most four of atoms of $J$ satisfies it automatically.

[45] Faraut and Korányi, *Analysis on Symmetric Cones*, note 16. Thm. III.1.2.

### 5.3 The four-point invariant

Where the invariant comes from: in a quaternionic Hilbert space choose unit vectors for four atoms with the three inner products with the first vector real and positive. The other three inner products are quaternions $g_1, g_2, g_3$, and the thirteen traces determine their real parts and the Gram matrix and triple product of their imaginary parts $u_1, u_2, u_3 \in \mathbb{R}^3$. In $\mathbb{R}^3$ the squared triple product equals the Gram determinant, which gives a polynomial in the traces vanishing on quaternionic quadruples. For octonions the $u_k$ lie in $\mathbb{R}^7$, where the Gram determinant exceeds the squared "triple product" $\langle u_1 \times u_3, u_2 \rangle$ by a quarter of the squared associator (Harvey and Lawson[46]). The polynomial measures that excess.

The invariant depends on the double-pair traces only through $\omega_1, \omega_2, \omega_3$, and in the variables $(o, R, \omega)$ it has sixteen terms:

$$\begin{aligned}\mathcal{A} = {} & o_{AB}o_{AC}o_{AD}o_{BC}o_{BD}o_{CD} - \big(R_{ABC}^2\, o_{AD}o_{BD}o_{CD} + R_{ABD}^2\, o_{AC}o_{BC}o_{CD} \\ & + R_{ACD}^2\, o_{AB}o_{BC}o_{BD} + R_{BCD}^2\, o_{AB}o_{AC}o_{AD}\big) \\ & - 4R_{ABC}R_{ABD}R_{ACD}R_{BCD} - \big(o_{AB}o_{CD}\,\omega_1^2 + o_{AC}o_{BD}\,\omega_2^2 + o_{AD}o_{BC}\,\omega_3^2\big) \\ & + 2\omega_1\big(R_{ABC}R_{ABD}\,o_{CD} + R_{ACD}R_{BCD}\,o_{AB}\big) \\ & + 2\omega_2\big(R_{ABC}R_{ACD}\,o_{BD} + R_{ABD}R_{BCD}\,o_{AC}\big) \\ & + 2\omega_3\big(R_{ABC}R_{BCD}\,o_{AD} + R_{ABD}R_{ACD}\,o_{BC}\big) - 2\,\omega_1\omega_2\omega_3.\end{aligned} \tag{5.4}$$

This polynomial is the definition of $\mathcal{A}$ for all atoms.

Where $o_{AB}o_{AC}o_{AD} > 0$, put $a = \sqrt{o_{AB}}$, $b = \sqrt{o_{AC}}$, $c = \sqrt{o_{AD}}$. Then

$$\mathcal{A} = a^2b^2c^2\big(\det G - s^2\big), \tag{5.5}$$

where the symmetric $3\times3$ matrix $G$ and the number $s$ are explicit rational functions of the traces and of $a, b, c$, displayed in (A.6). The polynomial is symmetric under all 24 relabellings of $A, B, C, D$, invariant under $F_4$, and of degree three in each atom; expanded in the thirteen traces it has 52 monomials (Appendix A.9).

In part (iv) of the next theorem, $i$, $j$ and $\ell$ are the octonion units of the Cayley–Dickson construction $\mathbb{O} = \mathbb{H} \oplus \mathbb{H}\ell$ with $\mathbb{H} = \operatorname{span}_{\mathbb{R}}\{1, i, j, ij\}$ (Appendix A.2).

**Theorem 5.5** (Four-point criterion)**.** *Let $p_A, p_B, p_C, p_D$ be atoms of $J$ and $\mathcal{A}$ the polynomial* (5.4) *in their thirteen traces. Then:*

*Formalized.* `thm_fourpoint`; the strict signs in (iv): `lemma_A_normal_nondegenerate`. Parts (i) and (iii): `AlbertChallenge.fourPoint_criterion`.

*(i)* $\mathcal{A} \geq 0$*;*

*(ii)* $\mathcal{A} = 0$ *if and only if the four atoms lie in a common quaternionic subalgebra of $J$;*

*(iii)* $\mathcal{A} > 0$ *if and only if* $\operatorname{Jord}\langle e, p_A, p_B, p_C, p_D\rangle = J$*;*

*(iv) if no two of the atoms coincide or are orthogonal, then there is $g \in F_4$ with $gp_i = v_i v_i^{\dagger}$ for the unit vectors*

$$v_A = (1,0,0),\quad v_B = (B_0, B_1, 0),\quad v_C = (C_0,\ C_1 + C_2 i,\ C_3),\quad v_D = (D_0, X, Y), \tag{5.6}$$

[46] R. Harvey and H. B. Lawson, Jr., Calibrated geometries, *Acta Math.* **148** (1982) 47–157. doi:10.1007/BF02392726. Thm. IV.1.6, p. 114.

*with* $B_0, B_1, C_0, D_0 > 0$, $C_2, C_3 \geq 0$, $C_1 \in \mathbb{R}$, $X = x_0 + x_1 i + x_2 j$ *and* $Y = y_0 + y_1 i + y_2 j + y_3\, ij + y_4 \ell$ *($x_k, y_k$ real, $x_2, y_4 \geq 0$), and then*

$$\mathcal{A} = \tfrac{1}{4}\, B_0^2 B_1^4\, C_0^2 C_2^2 C_3^2\, D_0^2\, \big|[i, X, Y]\big|^2 = B_0^2 B_1^4\, C_0^2 C_2^2 C_3^2\, D_0^2\, x_2^2 y_4^2. \tag{5.7}$$

*Here* $[x, y, z] = (xy)z - x(yz)$ *is the associator.*

That some four atoms generate $J$ was known (Section 9.6).

*Proof. Normal form.* Lemma A.4 moves the atoms to (5.6) using the stabiliser chain $F_4 \supset \mathrm{Spin}(9) \supset \mathrm{Spin}(8) \supset \mathrm{Spin}(7) \supset G_2$; the first three vectors lie in $\mathbb{C}_i^3$, which is Proposition 5.1 again.

*Traces become octonion data.* In the normal form put

$$g_1 = B_0 C_0 + B_1 z, \qquad g_2 = B_0 D_0 + B_1 X, \qquad g_3 = C_0 D_0 + \bar{z} X + C_3 Y, \qquad z = C_1 + C_2 i,$$

the octonionic analogues of the three edge amplitudes not touching $A$, and $u_k = \operatorname{Im} g_k$. Lemma A.5 writes the thirteen traces in the $g_k$ (A.5), without reassociating, and shows that $G$ is the Gram matrix $(\langle u_k, u_l \rangle)$ and $s = \langle u_1 \times u_3, u_2 \rangle$, where $u \times v = \operatorname{Im}(uv)$.

*The decisive step.* For imaginary octonions the Harvey–Lawson identity (A.1) gives $\det G - s^2 = \frac{1}{4}|[u_1, u_3, u_2]|^2$. The associator is trilinear and alternating, and it vanishes when one argument is real or when the three arguments lie in a subalgebra generated by two octonions (Artin's theorem). Since $z, \bar{z} \in \mathbb{C}_i$, this leaves $[g_1, g_3, g_2] = B_1^2 C_2 C_3 [i, Y, X]$, and (5.5) with $a = B_0$, $b = C_0$, $c = D_0$ gives the first form of (5.7). For $X, Y$ as in (5.6), $[i, X, Y] = x_2 [i, j, Y] = x_2 y_4 [i, j, \ell]$ and $|[i, j, \ell]|^2 = 4$, which is the second form. Lemma A.5 states this as a polynomial identity valid without normalising the vectors.

*(ii), "if".* Let $\psi : H_3(\mathbb{H}) \to J$ be a unital Jordan monomorphism whose image is a quaternionic subalgebra containing the atoms. It maps Jordan frames to Jordan frames of $J$, so it preserves the thirteen traces. Lemma A.7(a) brings the four atoms $\psi^{-1}(p_i)$ of the standard $H_3(\mathbb{H})$ to the shape (5.6) with $X, Y \in \mathbb{H}$, and Lemma A.5 gives $\mathcal{A} = \frac{1}{4}(\cdots)|[i, X, Y]|^2 = 0$, since $\mathbb{H}$ is associative.

*(i).* If two atoms coincide or are orthogonal, they lie in a quaternionic subalgebra (Lemma A.6) and $\mathcal{A} = 0$. Otherwise (5.7) applies.

*(ii), "only if".* Pairs that coincide or are orthogonal are covered by Lemma A.6. Otherwise $\mathcal{A} = 0$ means $C_2 C_3 x_2 y_4 = 0$ in (5.7), and Lemma A.7(b) shows that in each of these four cases the atoms lie in a quaternionic subalgebra.

*(iii).* If $\mathcal{A} > 0$, no pair coincides or is orthogonal, every factor in (5.7) is nonzero, and Lemma A.8 shows that the atoms generate $J$. Conversely, a quaternionic subalgebra is proper, so atoms that generate $J$ have $\mathcal{A} \neq 0$ by (ii), hence $\mathcal{A} > 0$ by (i). □

**The fourth atom in the complex subalgebra of the other three.** For atoms $p_1, p_2, p_3$ define the *triangle factor*

$$\Delta_W(p_1, p_2, p_3) = N(p_1 + p_2 + p_3)\left(o_{12} o_{23} o_{31} - R_{123}^2\right),$$
$$N(p_1 + p_2 + p_3) = 1 - o_{12} - o_{13} - o_{23} + 2R_{123}.$$

For atoms of a qutrit, $p_k = |a_k\rangle\langle a_k|$, the first factor is $|\det(a_1, a_2, a_3)|^2$ and the second is $o_{12}o_{23}o_{31}\sin^2\phi$, with $\phi$ the Bargmann phase of the triangle; so $\Delta_W \neq 0$ says that the three vectors are linearly independent and the triangle carries a non-real phase.

**Proposition 5.6** (The invariant as a determinant in a complex subalgebra)**.** *Let $p_1, p_2, p_3$ be atoms of $J$ with $\Delta_W(p_1, p_2, p_3) \neq 0$. Then $W = \mathrm{Jord}\langle e, p_1, p_2, p_3\rangle$ is a complex subalgebra, and for every atom $p_4$*

$$\mathcal{A}(p_1, p_2, p_3, p_4) = \Delta_W(p_1, p_2, p_3) \cdot N(P_W p_4),$$

*where $P_W$ is the trace-orthogonal projection of $J$ onto $W$ and $N$ restricted to $W$ is the determinant of $3\times3$ Hermitian matrices. If $\Delta_W(p_1, p_2, p_3) = 0$, then $\mathcal{A}(p_1, p_2, p_3, p_4) = 0$ for every atom $p_4$.*

*Formalized.* `prop_windowdet`, `prop_windowdet_zero`, for any complex subalgebra containing the three atoms; that it is the one they generate: `lemma_A_window_generated`.

*Proof.* Assume first $\Delta_W \neq 0$. By Lemma A.9, after relabelling $p_1, p_2, p_3$ and applying some $g \in F_4$ the atoms $(p_1, p_2, p_3, p_4)$ take the normal form (5.6) in the roles of $(A, B, C, D)$, with $gW = H_3(\mathbb{C}_i)$ and

$$\Delta_W = B_0^2 B_1^4 C_0^2 C_2^2 C_3^2. \tag{5.8}$$

Since $g$ is orthogonal and preserves $N$, the computation can be made in the normal form.

*The fourth atom.* Write $X = X_c + X_\perp$, $Y = Y_c + Y_\perp$ with $X_c, Y_c \in \mathbb{C}_i$ and $X_\perp = x_2 j$, $Y_\perp = y_2 j + y_3\, ij + y_4 \ell$ orthogonal to $\mathbb{C}_i$. The projection onto $H_3(\mathbb{C}_i)$ keeps the diagonal and replaces each off-diagonal entry by its $\mathbb{C}_i$-component. Since $\mathbb{C}_i\mathbb{C}_i^\perp + \mathbb{C}_i^\perp\mathbb{C}_i \subseteq \mathbb{C}_i^\perp$, the 23 entry $X\bar{Y}$ of $p_4 = v_D v_D^\dagger$ projects to $X_c\bar{Y}_c + \kappa$, where $\kappa = x_2(y_2 - y_3 i)$ is the $\mathbb{C}_i$-component of $X_\perp\bar{Y}_\perp$. Thus

$$P_W p_4 = uu^\dagger + E, \qquad u = (D_0, X_c, Y_c), \qquad E = \begin{pmatrix} 0 & 0 & 0 \\ 0 & x_2^2 & \kappa \\ 0 & \bar{\kappa} & |Y_\perp|^2 \end{pmatrix}.$$

Subtracting multiples of the first row, $\det(uu^\dagger + E) = D_0^2\big(x_2^2|Y_\perp|^2 - |\kappa|^2\big) = D_0^2 x_2^2 y_4^2$. Multiplying by (5.8) gives the right side of (5.7), which equals $\mathcal{A}$ by Lemma A.5.

If $\Delta_W = 0$, Lemma A.9(c) gives $\mathcal{A} = 0$. □

**Corollary 5.7** (Rank trichotomy)**.** *In the situation of Proposition 5.6 with $\Delta_W \neq 0$, $P_W p_4$ is a state of the complex subalgebra $W$, and*

*Formalized.* `cor_fp_rank`.

*(i)* $\operatorname{rank} P_W p_4 = 1$ *if and only if $p_4 \in W$;*

*(ii)* $\operatorname{rank} P_W p_4 \leq 2$ *if and only if $p_1, \ldots, p_4$ lie in a common quaternionic subalgebra (which then contains $W$);*

*(iii)* $\operatorname{rank} P_W p_4 = 3$ *if and only if $p_1, \ldots, p_4$ generate $J$.*

*Moreover $N(P_W p_4) \leq \frac{1}{27}$, with equality if and only if $P_W p_4 = e/3$.*

*Proof.* Positivity and trace one are Lemma A.2. In the normal form above, $u$ has nonzero first entry and $E$ is supported on the last two coordinates, so $\operatorname{rank} P_W p_4 = 1 + \operatorname{rank} E$. Now $E = 0$ iff $X, Y \in \mathbb{C}_i$ iff $p_4 \in H_3(\mathbb{C}_i)$, which is (i). The lower block of $E$ has determinant $x_2^2 y_4^2$, so $\operatorname{rank} E \le 1$ iff $N(P_W p_4) = 0$ iff $\mathcal{A} = 0$, which is (ii) by Theorem 5.5(ii); (iii) is the complement, by Theorem 5.5(iii). The last statement is the inequality between the arithmetic and geometric means of the eigenvalues of a positive matrix of trace one. □

The right-hand side of Proposition 5.6 is not manifestly symmetric; each choice of the fourth atom gives a different factorisation of the same number.

The quaternionic rank bound also has a converse for complex states. Every complex density matrix of rank two is the complex part of a quaternionic pure state, as shown by Asorey, Scolarici and Solombrino[47] and Masillo, Scolarici and Sozzo[48]. Indeed, write it as $xx^* + yy^*$ and take the quaternionic unit column $x + yj$, with $ji = -ij$. Its dyad $(x + yj)(x + yj)^*$ has complex part $xx^* + yy^*$. Taking $y = 0$ covers rank one.

## 5.4 The tetrahedron identity

Suppose a sequential projective representation, with projections $P_A, P_B, P_C, P_D$ and state $\varphi$, reproduces every word probability of the four tests of atoms $p_A, \dots, p_D$ from a state $\rho$. By Lemma 4.3 the argument passes to the real cyclic representation: a real Hilbert space $\mathcal{K}$, orthogonal projections $P_i$ on it (with the same notation), and a unit vector $\Omega$ with $\Pr(w) = \|P_w \Omega\|^2$. By Lemma 4.5, for every atom $p_a$ and every word $v$ the purity operator $P_a E_v P_a - c(v) P_a$ vanishes on $\mathcal{K}$, with $c(v) = \langle p_a, \theta(v) \rangle$: in $J$ a "yes" on an atom leaves a multiple of that atom.

**Lemma 5.8** (Relations forced by the statistics)**.** *In this situation, for all distinct labels, as operator identities on $\mathcal{K}$:*

$$\begin{aligned} &\textit{(R1)} \quad P_a P_b P_a = o_{ab} P_a, \\ &\textit{(R2)} \quad P_a (P_b P_c + P_c P_b) P_a = 2 R_{abc}\, P_a, \\ &\textit{(R3)} \quad P_a (P_b P_c P_d + P_d P_c P_b) P_a = 2 Q_{a;bcd}\, P_a. \end{aligned}$$

*Narrower in Lean.* `triangleDefect_primitive_decomposition`, `symmetricTripleDefect_primitive_decomposition`. Lean writes the differences of the two sides of (R2), (R3) as combinations of purity operators; the operator identities then follow from Lemma 4.5, a step not stated as a declaration.

*Proof.* (R1) is the purity relation with $v = (b)$. For (R2) and (R3) use (5.2) for operators, $L_k(Y) = \frac{1}{2}(P_k Y + Y P_k) = \frac{1}{2}\big(Y + P_k Y P_k - (1 - P_k) Y (1 - P_k)\big)$, and the identities, valid in any associative algebra,

$$P_k P_l + P_l P_k = 2 L_k(P_l), \qquad P_k P_l P_m + P_m P_l P_k = 2\big[L_k L_l(P_m) + L_m L_k(P_l) - L_l L_k(P_m)\big].$$

Expanding each $L$ writes both left sides as signed combinations of sequential products $E_v$ with $v$ of length at most two, resp. three. The same formal combinations of Lüders updates applied to $e$ in $J$ give $2p_k \circ p_l$ and $2[p_k \circ (p_l \circ p_m) + p_m \circ (p_k \circ$

[47] M. Asorey, G. Scolarici and L. Solombrino, Complex projections of completely positive quaternionic maps, *Theoret. and Math. Phys.* **151** (2007) 735–743. Math-Net.Ru. Prop. 4.

[48] F. Masillo, G. Scolarici and S. Sozzo, Proper versus improper mixtures: Toward a quaternionic quantum mechanics, *Theoret. and Math. Phys.* **160** (2009) 1006–1013. arXiv:0901.0795. Props. 2–3.

$p_l) - p_l \circ (p_k \circ p_m)]$, by (5.2) in $J$. Sandwiching by $P_a$ and applying purity term by term replaces each $P_a E_v P_a$ by $c(v)P_a$, and the sum of the $c(v)$ is the trace of the Jordan value against $p_a$, namely $2R_{akl}$, resp. $2Q_{a;klm}$ by (5.1). □

Two subspaces with projections $P, Q$ satisfying $PQP = \lambda P$ and $QPQ = \lambda Q$ are *isoclinic*: all their principal angles are equal, with $\cos^2 = \lambda$ (Wong[49], Wolf[50], Hoggar[51]); so (R1) is the isoclinic relation, and a representation of four atoms is a configuration of four pairwise isoclinic subspaces with parameters $o_{ab}$, subject to (R2) and (R3) (for algebras of projections with such relations see Section 9.4).

**Holonomies at a vertex.** Fix a base vertex $j$ and order the other labels as $k, l, m$; assume all six transition probabilities are positive. The three *triangle holonomies* at $j$ are

$$h_1 = \frac{P_j P_k P_l P_j}{\nu_1}, \quad h_2 = \frac{P_j P_l P_m P_j}{\nu_2}, \quad h_3 = \frac{P_j P_m P_k P_j}{\nu_3},$$
$$\nu_1 = \sqrt{o_{jk}o_{kl}o_{lj}},\ \nu_2 = \sqrt{o_{jl}o_{lm}o_{mj}},\ \nu_3 = \sqrt{o_{jm}o_{mk}o_{kj}},$$

operators on the corner $P_j\mathcal{K}$, split as $h_a = H_a + N_a$ into Hermitian part $H_a = \frac{1}{2}(h_a + h_a^*)$ and skew part $N_a = \frac{1}{2}(h_a - h_a^*)$.

On a qutrit with $P_i = |a_i\rangle\langle a_i|$, $h_1$ is multiplication by the Bargmann phase factor $\langle a_j|a_k\rangle\langle a_k|a_l\rangle\langle a_l|a_j\rangle/\nu_1$; in general it is the operator (non-Abelian) version, the holonomy of the geodesic triangle $jkl$. The shape invariants and double-pair traces are the real parts of the Bargmann invariants of three and four pure states[52], whose relations for complex vectors are studied by Oszmaniec, Brod and Galvão[53].

The relations attach to these operators the cosines of the Bargmann phases of the four triangles,

$$c_1 = \frac{R_{jkl}}{\nu_1}, \quad c_2 = \frac{R_{jlm}}{\nu_2}, \quad c_3 = \frac{R_{jmk}}{\nu_3}, \quad c_4 = \frac{R_{klm}}{\sqrt{o_{kl}o_{lm}o_{mk}}},$$

the three numbers

$$K_{12} = \frac{Q_{j;klm}}{\sqrt{o_{jk}o_{kl}o_{lm}o_{mj}}}, \quad K_{23} = \frac{Q_{j;lmk}}{\sqrt{o_{jl}o_{lm}o_{mk}o_{kj}}}, \quad K_{31} = \frac{Q_{j;mkl}}{\sqrt{o_{jm}o_{mk}o_{kl}o_{lj}}},$$

and the symmetric matrix $M$ with $M_{aa} = 1 - c_a^2$ and $M_{ab} = c_a c_b - K_{ab}$ $(a \neq b)$. Put

$$S = c_1c_2c_3 - c_1M_{23} - c_2M_{13} - c_3M_{12}, \qquad \mathcal{D} = \det M - (c_4 - S)^2. \tag{5.9}$$

This is the tetrahedron statistic, a function of the thirteen traces and hence of word probabilities of length at most four.

[49] Wong, *Isoclinic n-planes*, note 11.

[50] Wolf, "Elliptic spaces in Grassmann manifolds", note 11.

[51] S. G. Hoggar, New sets of equi-isoclinic $n$-planes from old, *Proc. Edinburgh Math. Soc.* **20** (1977) 287–291. doi:10.1017/S0013091500026511. Prop. 1.2(b).

[52] Bargmann, "Note on Wigner's theorem", note 2.

[53] M. Oszmaniec, D. J. Brod and E. F. Galvão, Measuring relational information between quantum states, and applications, *New J. Phys.* **26** (2024) 013053. doi:10.1088/1367-2630/ad1a27. arXiv:2109.10006.

**Theorem 5.9** (Tetrahedron identity)**.** *Let $P_j, P_k, P_l, P_m$ be orthogonal projections on a real Hilbert space $\mathcal{K}$ of any dimension, with $P_j \neq 0$. Let $o_{ab} > 0$, $R_{abc}$ and $Q_{j;\cdot}$ be real numbers such that, as operator identities on $\mathcal{K}$, (R1) holds for all pairs, (R2) holds at the bases $j$ and $k$, and (R3) holds at the base $j$. Define $\mathcal{D}$ from these numbers by* (5.9)*. Then $\mathcal{D} = 0$.*

*Narrower in Lean.* `tetrahedronEngine`. Lean proves the algebraic form, Theorem F.3; the passage to $\mathcal{D} = 0$ on a Hilbert space is not stated.

The precedent is Wolf's Lemma 7 in[54], discussed below. No division by $\sqrt{1 - c_a^2}$ occurs, so real triangles ($c_a = \pm 1$) and degenerate $M$ are covered.

*Proof.* The bounded operators on $\mathcal{K}$ form an associative real algebra in which the $P_a$ are idempotents satisfying the hypotheses of Theorem F.3. So $\Delta_j P_j = 0$, and $\Delta_j = 0$ because $P_j \neq 0$. In (F.4) the products of transition probabilities are $\kappa_a = \nu_a^2$, and with $D = \operatorname{diag}(\nu_1, \nu_2, \nu_3)$ the definitions give $\mathcal{B} = -DMD$ and $\sigma = \nu_1\nu_2\nu_3(c_4 - S)$; for instance $b_{12} = \nu_1\nu_2(K_{12} - c_1c_2)$ because $\nu_1\nu_2 = o_{jl}\sqrt{o_{jk}o_{kl}o_{lm}o_{mj}}$. Hence $0 = \Delta_j = -\det\mathcal{B} - \sigma^2 = (\nu_1\nu_2\nu_3)^2\,\mathcal{D}$, and $\mathcal{D} = 0$ since every $o_{ab} > 0$. □

On the corner $P_j\mathcal{K}$, with 1 denoting $P_j$, the steps of the proof of Theorem F.3, divided by the $\nu_a$, are statements about the holonomies. Step (b) says that each $h_a$ is unitary and step (a) that its Hermitian part is $H_a = c_a$, so $N_a^2 = -M_{aa}$, and step (d) gives the Clifford relations

$$N_aN_b + N_bN_a = -2M_{ab} \qquad (a, b = 1, 2, 3). \tag{5.10}$$

After the contractions $P_lP_jP_l = o_{jl}P_l$ and $P_mP_jP_m = o_{jm}P_m$ the product $h_1h_2h_3$ is a multiple of $P_jP_kP_lP_mP_kP_j$, the holonomy of the far face, and step (e) says

$$\tfrac{1}{2}\big(h_1h_2h_3 + (h_1h_2h_3)^*\big) = c_4. \tag{5.11}$$

Expanding $h_1h_2h_3 = (c_1 + N_1)(c_2 + N_2)(c_3 + N_3)$, the left side is $S + \Gamma$ with the volume element $\Gamma = \frac{1}{2}(N_1N_2N_3 - N_3N_2N_1)$ of the Clifford relations, and step (f), the one place where associativity is used, is

$$\Gamma = N_1N_2N_3 + M_{23}N_1 - M_{13}N_2 + M_{12}N_3, \qquad \Gamma^2 = \det M \tag{5.12}$$

(Lemma A.10). So $\Gamma = c_4 - S$ and $(c_4 - S)^2 = \det M$.

The argument never treats $\Gamma$ as a sign $\pm 1$, chooses no Clifford module and does not normalise $M$, so a direct sum of models with different Clifford sectors gives no escape. It needs the relations only on finitely many vectors (Section 6.3).

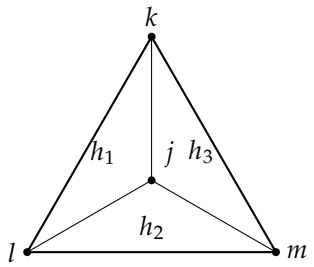
**Proposition 5.10** (The tetrahedron statistic is the invariant)**.** *For every choice of base $j$ and order $k, l, m$,*

$$\mathcal{D} \cdot o_{AB}\,o_{AC}\,o_{AD}\,o_{BC}\,o_{BD}\,o_{CD} = \mathcal{A},$$

*as an identity of rational functions in the thirteen traces regarded as independent variables with positive transition probabilities. In particular $\mathcal{D}$ does not depend on the base or the order.*

*Formalized.* `prop_Dacal_poly`; `tetrahedronDiscriminant_eq_overlap_mul_fourPointInvariant` is a variant normalised at one vertex.

[54] Wolf, "Elliptic spaces in Grassmann manifolds", note 11.




far face $klm$:
cosine $c_4$

Figure 1: Schematic, drawn flat with the vertex $j$ in the middle. The holonomies $h_1, h_2, h_3$ go around the triangles $jkl$, $jlm$, $jmk$; their product, after two contractions by (R1), is the holonomy around the far face $klm$. In a sequential projective representation this closes up ($\mathcal{D} = 0$); for four atoms of $J$ in general position it fails by a quarter of a squared associator.

*Proof.* By the symmetry of $\mathcal{A}$ it suffices to take $j = A$, $(k, l, m) = (B, C, D)$. Put $d = \sqrt{o_{BC}}$, $e' = \sqrt{o_{BD}}$, $f = \sqrt{o_{CD}}$, let $G, s$ be as in (5.5), and let $G'$ be $G$ with indices in the order $(1, 3, 2)$ (edges $BC$, $CD$, $BD$, matching $h_1, h_2, h_3$). Substituting the definitions gives

$$M = T\,G'\,T, \qquad T = \operatorname{diag}\Big(\frac{1}{d}, \frac{1}{f}, -\frac{1}{e'}\Big), \qquad c_4 - S = \frac{s}{d\,e'\,f}.$$

Hence $\mathcal{D} = (\det G - s^2)/(de'f)^2$; multiplying by $\prod o_{ij} = a^2b^2c^2d^2e'^2f^2$ and using (5.5) gives $\mathcal{A}$. No relation among the traces is used, in particular not (5.3). □

**Corollary 5.11** (No sequential projective representation)**.** *Let $p_A, p_B, p_C, p_D$ be atoms of $J$ with $\mathcal{A} > 0$. The model class is (I). No sequential projective representation reproduces all word probabilities of the four tests from $e/3$.*

*Formalized.* `LengthSeven.Uniform.noideal_algebraic`, `LengthSeven.Uniform.noideal_hilbert`, `LengthSeven.Uniform.noideal_rightQuaternionic`, from every state; also `FiveTest.cor_fp_noideal`.

*Proof.* If $\mathcal{A} > 0$, no two atoms coincide or are orthogonal (Theorem 5.5), so all $o_{ij} \in (0, 1)$. In a reproducing representation Lemma 5.8 supplies the hypotheses of Theorem 5.9, with $P_j \neq 0$ on $\mathcal{K}$ because $\|P_j\Omega\|^2 = \Pr(j) = \frac{1}{3}$. So $\mathcal{D} = 0$, and Proposition 5.10 gives $\mathcal{A} = \mathcal{D} \prod o_{ij} = 0$, a contradiction. □

This is the four-atom case of Theorem 4.1, which covers every state, by a different route. If $\mathcal{A} = 0$, the atoms lie in a quaternionic subalgebra and the standard embedding $H_3(\mathbb{H}) \to H_6(\mathbb{C})$ gives a representation on $\mathbb{C}^6$ with rank-two projections.

*Formalized.* $\mathcal{A} = 0$: `fourPoint_zero_c6_representation`, with fixed rank-two projections; with only the space $\mathbb{C}^6$ recorded, `AlbertChallenge.zero_invariant_model`.

**Octonionic holonomies.** In the normal form (5.6) put $h_1 = g_1/|g_1|$, $h_2 = g_3/|g_3|$, $h_3 = \bar{g}_2/|g_2|$: three unit octonions, the phases of the triangles $ABC$, $ACD$, $ADB$ read with all edges at $A$ real and positive.

By (A.5), $c_a = \operatorname{Re} h_a$, $M$ is the Gram matrix of $v_a = \operatorname{Im} h_a$, and $c_4 = \operatorname{Re}((h_1h_2)h_3)$. Imaginary octonions satisfy the Clifford relations $v_av_b + v_bv_a = -2\langle v_a, v_b\rangle$, like the $N_a$, so the octonionic phases obey everything a sequential projective representation demands of the holonomies, including the scalar relation (5.11). In any associative model these relations fix $\Gamma^2 = \det M$ by regrouping products. For octonions,

$\mathrm{Re}((h_1h_2)h_3) - S = \mathrm{Re}((v_1v_2)v_3) = -\langle v_1 \times v_2, v_3\rangle$, and the Harvey–Lawson identity gives

$$\mathcal{D} = \det M - \langle v_1 \times v_2, v_3\rangle^2 = \tfrac{1}{4}\big|[v_1, v_2, v_3]\big|^2 = \tfrac{1}{4}\big|[h_1, h_2, h_3]\big|^2 :$$

the octonions miss by a squared associator. The tetrahedron identity holds exactly when the three phase axes lie in an associative 3-plane, the imaginary part of a quaternion subalgebra.

The same skeleton appears in Wolf's study of totally geodesic families of pairwise isoclinic subspaces[55]. When the projective lines of such a family have real dimension four, the family supplies three anticommuting unitaries satisfying the Hurwitz equations, and its closure under the Lie triple product forces $A_1A_2A_3 = \pm 1$ (pp. 454–455). The Hurwitz equations alone do not force this, since the sum of the two irreducible Clifford modules with opposite volume elements satisfies them with $A_1A_2A_3$ not scalar. The same argument rules out lines of dimension greater than four, which excludes the Cayley plane. Wolf assumes a whole totally geodesic family in a finite-dimensional Grassmannian over an associative field; here the input is four points in an arbitrary representation, the parameters $o_{ab}$ vary, $M$ may be degenerate, and the failure is measured by $\mathcal{A}$. Neither result implies the other.

## 5.5 How large can the invariant be?

**Proposition 5.12** (Maximum)**.** *For all atoms $p_A, p_B, p_C, p_D$ of $J$, $0 \le \mathcal{A} \le \frac{4}{19683} = \frac{4}{3^9}$. Equality holds on exactly one $F_4$-orbit of ordered quadruples, represented by the atoms $v_iv_i^\dagger$ with*

$$v_A = (1,0,0), \quad v_B = \tfrac{1}{\sqrt{3}}(1,\sqrt{2},0), \quad v_C = \tfrac{1}{\sqrt{3}}(1,i,1), \quad v_D = \tfrac{1}{\sqrt{3}}(1,j,\ell). \tag{5.13}$$

*There all transition probabilities are $\frac{1}{3}$, all shape invariants $\frac{1}{9}$, all double-pair traces $\frac{1}{27}$, and $\mathcal{D} = \frac{4}{27}$.*

*Formalized.* `prop_max`, `maximizerConfig_traces`, `fourPointInvariant_maximizerConfig`. The bound and the value at the maximiser: `AlbertChallenge.quadrupleM_spec`.

*Proof.* Configurations with a coinciding or orthogonal pair have $\mathcal{A} = 0$. Otherwise, in the normal form, $\mathcal{A} = (B_0^2B_1^4)(C_0^2C_2^2C_3^2)(D_0^2x_2^2y_4^2)$ with unit vectors.

The first factor is $u(1-u)^2$ with $u = B_0^2$, at most $\frac{4}{27}$, with equality only at $u = \frac{1}{3}$. The second and third are products of three nonnegative numbers with sum at most one, so at most $\frac{1}{27}$, with equality only when the three numbers are $\frac{1}{3}$ and all other coordinates vanish. This gives the bound, and equality forces the coordinates of (5.13); since every maximiser has no coinciding or orthogonal pair, the normal form applies to it, so the maximisers form one orbit. The traces are a direct computation, and $\mathcal{D} = \mathcal{A}/\prod o_{ij} = \frac{4}{3^9} \cdot 3^6$. □

Roelands and Tiersma[56] show by direct computation that a closely related quadruple of atoms generates $J$.

[55] Ibid. Prop. 1, Lemma 7, Thm. 2.

[56] Roelands and Tiersma, "On the Shirshov–Cohn theorem for JB-algebras", note 9. Lemma 6.5.

**Lemma 5.13** (Bound for the triangle factor)**.** *For any three atoms of $J$, $\Delta_W(p_1,p_2,p_3) \le \frac{4}{729}$, with equality if and only if $o_{12}=o_{13}=o_{23}=\frac{1}{3}$ and $R_{123}=\frac{1}{9}$.*

*Formalized.* `lem_fp_DeltaWmax`.

*Proof.* Let $\Delta_W \neq 0$. Lemma A.9(b), applied with $p_4 = p_1$, gives $0 < o_{ij} < 1$, and Lemma A.9(a) (whose computation uses only the first three vectors) gives, in the normal form of Lemma A.3 with the dyad vectors (A.4), $\Delta_W = (B_0^2 B_1^4)(C_0^2 C_2^2 C_3^2)$ with $B_0^2 + B_1^2 = 1$ and $C_0^2 + C_1^2 + C_2^2 + C_3^2 = 1$. As in the proof of Proposition 5.12, the two factors are at most $\frac{4}{27}$ and $\frac{1}{27}$, with equality only if $B_0^2 = \frac{1}{3}$, $C_1 = 0$ and $C_0^2 = C_2^2 = C_3^2 = \frac{1}{3}$. At such a point

$$o_{12} = B_0^2 = \tfrac{1}{3}, \quad o_{13} = C_0^2 = \tfrac{1}{3}, \quad o_{23} = |B_0 C_0 + B_1 C_2 i|^2 = \tfrac{1}{3},$$
$$R_{123} = B_0 C_0 (B_0 C_0 + B_1 C_1) = \tfrac{1}{9}.$$

Conversely, these values give $\Delta_W = \frac{2}{9}\cdot\left(\frac{1}{27}-\frac{1}{81}\right) = \frac{4}{729}$ directly from the definition. □

**Corollary 5.14** (At a maximiser every atom looks maximally mixed to the other three)**.** *Let $p_A, p_B, p_C, p_D$ be atoms of $J$ with $\mathcal{A} = \frac{4}{19683}$. For each $k$, the other three atoms generate a complex subalgebra $W_k$, and $P_{W_k} p_k = \frac{1}{3}e$. Consequently, for every word $w$ in the three tests other than $k$, the probability of $w$ from the pure state $p_k$ equals its probability from the maximally mixed state $e/3$.*

*Formalized.* `cor_fp_maxmixed`.

*Proof.* $\mathcal{A}$ is symmetric, so Proposition 5.6 applies with $p_k$ in the role of $p_4$; as $\mathcal{A} \neq 0$, $\Delta_W \neq 0$ for the other three, $W_k$ is a complex subalgebra, and $\mathcal{A} = \Delta_W \cdot N(P_{W_k} p_k)$. By Lemma 5.13 and Corollary 5.7, $\Delta_W \le \frac{4}{729}$ and $N(P_{W_k} p_k) \le \frac{1}{27}$. The product of the bounds is $\frac{4}{19683} = \mathcal{A}$, so both are attained, and $N(P_{W_k} p_k) = \frac{1}{27}$ means $P_{W_k} p_k = \frac{1}{3}e$. The last sentence follows from Lemma A.2. □

Corollary 5.14 has a counterpart for a random atom. If $p$ is uniformly distributed on the atoms of $J$ (its law is the $F_4$-invariant probability measure) and $W$ is a complex subalgebra, then $P_W p$, transported to $H_3(\mathbb{C})$, has the law of the reduced state on $\mathbb{C}^3$ of a uniformly random pure state of $\mathbb{C}^3 \otimes \mathbb{C}^4$, the induced measure of Życzkowski and Sommers[57]; the statement and its proof are in the companion paper (section "What one complex subalgebra sees" and appendix "Proof of the law of a pure state seen in a complex subalgebra"). This law and Corollary 5.14 say that exceptional purity looks like ordinary mixedness to an observer who has only three tests. Both compare statistics only: no tensor factor, environment or purification of $J$ is asserted.

## 5.6 An invariant of atoms is not a test of a table

Two questions should be kept apart. For a specified quadruple of Albert atoms, $\mathcal{A}$ is computed from the thirteen traces, and $\mathcal{A} > 0$ says that the atoms generate $J$ (Theorem 5.5).

For a table of word probabilities, the question is whether some sequential projective representation reproduces it. The thirteen traces are read from words of

[57] K. Życzkowski and H.-J. Sommers, Induced measures in the space of mixed quantum states, *J. Phys. A* **34** (2001) 7111–7125. doi:10.1088/0305-4470/34/35/335. arXiv:quant-ph/0012101.

length at most four, so a representation has them too, and with them a value of $\mathcal{D}$. But a representation can reproduce all thirteen traces of the generating quadruple $\mathsf{S}$ (Section 6, paragraph "Two configurations"), with $\mathcal{D} = \frac{1}{12} > 0$ (Proposition 6.16), and the companion paper (section "The left-multiplication model") gives a real representation that reproduces all thirteen traces of any generating quadruple. The tetrahedron identity forces $\mathcal{D} = 0$ only through the purity relations, which are statements about longer words. Corollary 5.11 uses them on all words.

# 6 Finite certificates

Four tests of atoms with $\mathcal{A} > 0$ have no sequential projective representation that matches every word (Section 5). Finitely many words already suffice, and from $e/3$ the least length is seven.

Throughout, the model class is class (I) of Definition 3.4. Words are written in chronological order with $i$ for "yes" and $\bar{\imath}$ for "no" on test $i$; a word is *reduced* if no two consecutive letters belong to the same test. In every sequential projective representation $Q_a Q_a = Q_a$ and $P_i(1 - P_i) = 0$, so the probability of any word is that of a reduced word or zero; only reduced words are therefore listed. The notation below uses $E_w$, the word operators $P_w$ and the real cyclic representation $(\pi, \Omega)$ of Lemma 4.3, the purity operators $z_q(v)$ of Lemma 4.5, the relations (R1)–(R3) of Lemma 5.8, and the holonomies $h_a = H_a + N_a$, the numbers $c_a$, $S$, the matrix $M$ and the volume element $\Gamma = \frac{1}{2}(N_1 N_2 N_3 - N_3 N_2 N_1)$ of Section 5.4.

For four atoms with $\mathcal{A} > 0$, let the *minimum separating word length* $L_{\min}$ be the least $L$ such that no sequential projective representation reproduces the octonionic probabilities from $e/3$ of all words of length at most $L$.

**Theorem 6.1** (Words of length seven suffice for every generating quadruple)**.** *Let $p_A, p_B, p_C, p_D$ be atoms of $J$ with $\mathcal{A} > 0$; the model class is (I).*

*(a) No sequential projective representation, over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$, of any dimension, with projections of any rank and any state, reproduces from $e/3$ the probability of every word of length at most seven that begins with "yes" on $A$; none reproduces from $p_A$ the probability of every word of length at most six. Hence $L_{\min} = 7$.*

*(b) There is $\delta > 0$, depending on the quadruple, such that every sequential projective representation differs from the octonionic table from $e/3$ by at least $\delta$ on some word of length at most seven that begins with "yes" on $A$.*

The proof is given in Section 6.2, after the lemmas it uses. The lower bound $L_{\min} \geq 7$ is the next proposition, where $\iota_{\mathbb{H}} : H_3(\mathbb{H}) \to H_6(\mathbb{C})$ denotes the standard embedding.

*Narrower in Lean.* (a) `AlbertChallenge.seven_steps_exclude`, `LengthSeven.Uniform.uniform_lengthSeven_e3_algebraic`, `LengthSeven.Uniform.uniform_lengthSix_pA_algebraic` and their Hilbert and quaternionic versions; $L_{\min} = 7$: `AlbertChallenge.least_separating_length`, `MatchingModels.lMin_eq_seven`. The minimum is taken over Hilbert spaces in Lean's lowest universe; the exclusions hold in every universe. (b) `Ends.uniform_b`; also a margin from $p_A$ at length six, `Ends.uniform_b_pA`. Below Lemma 6.5 the formal proof of (a) fixes the Gram table directly and does not use Lemma 6.6 (Appendix G).

**Proposition 6.2** (Six-step models for every generating quadruple)**.** *Let $p_A, p_B, p_C, p_D$ be atoms of $J$ with $\mathcal{A} > 0$. There are atoms $p_i^+$ and $p_i^-$ ($i = A, \dots, D$) of $H_3(\mathbb{H})$ and $\lambda \in (0,1)$ such that the projections $\iota_{\mathbb{H}}(p_i^+) \oplus \iota_{\mathbb{H}}(p_i^-)$ on $\mathbb{C}^{12}$ with the density matrix $\lambda I_6/6 \oplus (1-\lambda) I_6/6$, a mixture of two quaternionic quadruples, form a sequential projective representation that reproduces from $e/3$ every word of length at most six. Consequently, with Theorem 6.1, $L_{\min} = 7$ for every generating quadruple, and no linear equality valid in every sequential projective representation and supported on words of length at most six is violated by its table from $e/3$.*

*Narrower in Lean.* `MatchingModels.prop_cer_universalsix_witness`, `MatchingModels.prop_cer_universalsix_c12`, `AlbertChallenge.six_step_model`; the last sentence: `AlbertChallenge.least_separating_length`, `MatchingModels.lMin_eq_seven`, `MatchingModels.no_classI_linearEq_le_six`, and for inequalities `MatchingModels.no_classI_linearIneq_le_six`. The minimum is taken over Hilbert spaces in Lean's lowest universe; the exclusions hold in every universe.

*Proof.* Appendix C.6 constructs the two quaternionic quadruples, from the two ends of a chord of complex density matrices through the projection of the fourth atom onto the complex subalgebra of the other three. They have the transition probabilities and shape invariants of the Albert quadruple, and the same average double-pair traces. It proves that every word probability through length six is affine in those double-pair traces when the other data are fixed. The mixture therefore reproduces these words, so $L_{\min} \geq 7$; the upper bound is Theorem 6.1. The equality assertion follows because the matching model is itself a sequential projective representation. □

**Why six steps are affine.** Expand a word probability from $e/3$ with $U_p x = \langle p, x \rangle p$ for a "yes" and $U_{e-p} = \mathrm{id} - 2L_p + U_p$ for a "no". Each "yes" splits a term into trace factors, and the intervals of positions used by two factors share at most the splitting position. A factor that involves all four atoms within at most seven positions contains some atom only once; moving that atom to the outside of the trace pairing, and using the complex subalgebra generated by the other three, shows that the factor is affine in the double-pair traces. Two such factors would need at least $4 + 4 - 1 = 7$ positions. So through length six each term has at most one factor that depends on the double-pair traces, and it depends on them affinely (Appendix C.6).

The same theorem gives seven steps from every preparation, provided the first outcome may be a "yes" on any of the four tests.

**Corollary 6.3** (Seven steps from every state)**.** *Let $p_A, p_B, p_C, p_D$ be atoms of $J$ with $\mathcal{A} > 0$, and let $\rho$ be any state of $J$; the model class is (I). No sequential projective representation, over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$, of any dimension, with projections of any rank and any state, reproduces from $\rho$ the probability of every word of length at most seven whose first outcome is "yes". For $\rho = e/3$ the length seven cannot be replaced by six.*

*Formalized.* `BlockB.cor_cer_everystate`, `BlockB.everystate_lengthSeven_algebraic` and its Hilbert and quaternionic versions.

*Proof.* Let $y = p_A + p_B + p_C + p_D$. If $y$ were not invertible, an atom $r$ with $\langle r, y \rangle = 0$ would have $\langle r, p_i \rangle = 0$ for each $i$, so every $p_i$ would lie in $J_0(r)$ (Lemma 3.1) and the atoms would generate a subalgebra of the proper subalgebra $\mathbb{R}r \oplus J_0(r)$, against Theorem 5.5(iii). So $y$ is invertible and positive, $\langle \rho, y \rangle > 0$, and $a = \langle \rho, p_j \rangle > 0$ for some $j$. Suppose a representation with state $\varphi$ reproduces the words. Then $\varphi(P_j) = a$, and $\varphi_j(X) = a^{-1}\varphi(P_j X P_j)$ is again a state. For a word $v$ of length at most six, $E_{(j,v)} = P_j E_v P_j$, and in $J$ the "yes" on $j$ leaves $U_{p_j}\rho = a\, p_j$; hence $\varphi_j(E_v) = a^{-1} \Pr_\rho(j, v) = \Pr_{p_j}(v)$. So the same projections with the state $\varphi_j$ reproduce from $p_j$

every word of length at most six, which Theorem 6.1(a), with the labels exchanged so that $j$ plays the part of $A$, forbids. The last sentence is Proposition 6.2. □

From the maximally mixed state the first outcome may be fixed to be a "yes" on $A$; from a general state no single first test will do, since $\rho$ may give $p_A$ weight zero.

**Two configurations.** Let $(i, j, \ell)$ be a basic triple of unit imaginary octonions ($i \perp j$, and $\ell \perp 1, i, j, ij$; Appendix A.5). The *maximal configuration* M is the maximiser (5.13) of Proposition 5.12, the atoms $vv^\dagger/|v|^2$ of

$$v_A = (1, 0, 0), \quad v_B = (1, \sqrt{2}, 0), \quad v_C = (1, i, 1), \quad v_D = (1, j, \ell).$$

The configuration S replaces $v_B$ by $(1, 1, 0)$; all its vectors have rational octonion coordinates, so its traces are rational. Both generate $J$. For M all transition probabilities are $\frac{1}{3}$, all shape invariants $\frac{1}{9}$, all double-pair traces $\frac{1}{27}$, $\mathcal{D} = \frac{4}{27}$ and $\mathcal{A} = \frac{4}{19683}$. For S, $o_{AB} = \frac{1}{2}$ and the other transition probabilities are $\frac{1}{3}$, $(R_{ABC}, R_{ABD}, R_{ACD}, R_{BCD}) = (\frac{1}{6}, \frac{1}{6}, \frac{1}{9}, \frac{1}{9})$, $t_1 = t_2 = t_3 = \frac{1}{18}$, $\mathcal{D} = \frac{1}{12}$ and $\mathcal{A} = \frac{1}{5832}$. The traces of both are derived in the proof of Proposition 6.11, and the values of $\mathcal{A}$ follow from (5.4): for S, with $\omega_r = \frac{1}{18}$, its eight groups of terms are $\frac{1}{486}$, $-\frac{1}{486} - \frac{1}{729}$, $-\frac{1}{729}$, $-\frac{7}{5832}$, $\frac{5}{2916}$, $\frac{1}{729}$, $\frac{1}{729}$ and $-\frac{1}{2916}$, with sum $\frac{1}{5832}$, and $\mathcal{D} = \mathcal{A}/\prod o_{ab} = 486/5832$ (Proposition 5.10); for M see Proposition 5.12. Appendix C.4 explains why the configurations, and hence every number below, do not depend on the choice of octonion basis.

*Narrower in Lean.* S: `AlbertChallenge.quadrupleS_spec`, that its elements are atoms and $\mathcal{A} = \frac{1}{5832}$; its traces: `BlockB.isTracesS_fixedS`. M: `AlbertChallenge.quadrupleM_spec`, `BlockB.isTracesM_maximizerConfig`. The value of $\mathcal{D}$ for S is not stated in this form.

## 6.1 Finite lists

**The model side at length $L$.** The following lemma is the finite-level form of the moment-matrix characterisation of Budroni, Moroder, Kleinmann and Gühne[58]. Their appendix sketches the converse by reference to the proof of Theorem 8 of Navascués, Pironio and Acín[59], which treats the limit of all orders; the truncated statement needs only the short argument below. Let $\mathcal{W}_L$ be the set of reduced words of length at most $L$ in the letters $i, \bar{\imath}$ $(i = 1, \ldots, n)$, the empty word included. A *table on* $\mathcal{W}_L$ is a function $f : \mathcal{W}_L \to [0, 1]$ with $f(\varnothing) = 1$.

**Lemma 6.4** (Reconstruction at finite length)**.** *A table $f$ on $\mathcal{W}_L$ is the restriction of the table of a sequential projective representation if and only if there is a real positive semidefinite matrix $\mathcal{G}$, indexed by $\mathcal{W}_L$, with*

*(M1)* $\mathcal{G}_{w,w} = f(w)$ *for every* $w \in \mathcal{W}_L$*;*

*(M2)* $\mathcal{G}_{u,vi} + \mathcal{G}_{u,v\bar{\imath}} = \mathcal{G}_{u,v}$ *for all* $u \in \mathcal{W}_L$ *and all* $v \in \mathcal{W}_{L-1}$ *whose last letter is not a letter of test* $i$*;*

*(M3)* $\mathcal{G}_{u,v} = 0$ *whenever* $u$ *and* $v$ *end with the two different outcomes of the same test.*

*Different route.* `finiteGram_iff_finiteRealModel`, `AlgebraicProjectiveData.finiteGram`: necessity is proved from the positive functional of an algebraic model, not through the cyclic representation.

[58] Budroni et al., "Bounding temporal quantum correlations", note 3. eqs. (7)–(8) and App. A3.

[59] M. Navascués, S. Pironio and A. Acín, A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations, *New J. Phys.* **10** (2008) 073013. doi:10.1088/1367-2630/10/7/073013. arXiv:0803.4290.

*In that case there is a representation on a real Hilbert space of dimension* $\operatorname{rank}\mathcal{G}$ *with a vector state.*

The proof is in Appendix C.1.

The set of tables on $\mathcal{W}_L$ of sequential projective representations is therefore convex and compact: direct sums give convex combinations, and $|\mathcal{G}_{u,v}| \le 1$.

**The group algebra.** With $\Sigma_i = 2P_i - 1$ as in Section 4.3, in every representation $\Sigma_i$ is a self-adjoint unitary. Let $\mathbb{R}[\mathbb{Z}_2^{*n}]$ be the real group algebra of the free product $\mathbb{Z}_2^{*n} = \mathbb{Z}_2 * \cdots * \mathbb{Z}_2$ ($n$ factors) with generators $r_1, \dots, r_n$; its basis is the set of reduced words in the $r_i$, and $u \mapsto u^{\mathrm{rev}}$ (reversal) is its involution. A formal word combination $\Lambda = \sum_w \lambda_w [w]$ has the element $E(\Lambda) = \sum_w \lambda_w E_w$ of $\mathbb{R}[\mathbb{Z}_2^{*n}]$, with each $P_i$ written as $\frac{1}{2}(1 + r_i)$. On the octonionic side, $\Lambda$ has the effect $\theta(\Lambda) = \sum_w \lambda_w \theta(w) \in J$ of Section 4.2; no map from $\mathbb{R}[\mathbb{Z}_2^{*n}]$ to $J$ is needed.

## 6.2 Words of length seven suffice and length six does not

The proof of Theorem 6.1 builds, from a representation that matches the words, three real vectors whose Gram matrix the word probabilities force to have negative determinant. It runs in four steps, which this subsection takes in order. First, conditioning on "yes" to $A$ gives hypothesis (H) below, and the words force relations on short histories (Lemma 6.5). Second, these relations determine every walk value of length at most eight (an inner product $\langle \xi, P_{w_1} \cdots P_{w_n} \xi \rangle$, defined below) as a fixed polynomial in the traces (Lemma 6.6). Third, an explicit indefinite form, the comparison form, satisfies the same relations (Lemma 6.7), so the words fix the Gram matrix of twenty-eight short path vectors to be their Gram matrix under that form (Lemma 6.8). Fourth, in that form three combinations of them have Gram determinant $-4\hat{D}^3/\prod_x(1 - o_{Ax})$, where $\hat{D} = \mathcal{A}/(o_{AB}o_{AC}o_{AD})^3$ (Lemma 6.9); the proof of the theorem follows that lemma.

The rest of the subsection treats M and S by one explicit vector (Theorem 6.10), shows that six steps do not suffice for them (Proposition 6.11), and explains why the six-step mixture fails at seven (Remark 6.12, Proposition 6.13). On a first reading the proofs of Lemmas 6.6–6.9 can be skipped.

**The conditioned representation.** Suppose a sequential projective representation reproduces, from $e/3$, the probability of every word of length at most seven that begins with "yes" on $A$. Pass to its real cyclic representation (Lemma 4.3): a real Hilbert space $\mathcal{K}$, orthogonal projections on it, written again $P_A, \dots, P_D$, and a unit vector $\Omega$ with $\Pr(w) = \|P_w\Omega\|^2$. Since $\Pr(A) = \frac{1}{3} > 0$, the vector $\xi = P_A\Omega/\|P_A\Omega\|$ is a unit vector with $P_A\xi = \xi$, and since $U_{p_A}(e/3) = \frac{1}{3}p_A$,

$$\|P_u\xi\|^2 = 3\Pr_{e/3}(A, u) = \Pr_{p_A}(u) \qquad \text{for every word } u \text{ of length at most six.} \quad (6.1)$$

A representation that reproduces from the pure state $p_A$ every word of length at most six gives (6.1) directly, with $\xi = \Omega$: there $\|(1 - P_A)\Omega\|^2 = 1 - \Pr(A) = 0$. Below, *hypothesis (H)* means: a real Hilbert space with orthogonal projections $P_A, \dots, P_D$

and a unit vector $\xi = P_A\xi$ satisfying (6.1). Only this direction of the conditioning is used.

All inner products are real. A *walk* is a word $w = w_1 \cdots w_n$ in the letters $A, B, C, D$, standing for the operator $P_{w_1} \cdots P_{w_n}$, and

$$L(w) = \langle \xi, P_{w_1} \cdots P_{w_n} \xi \rangle.$$

$L$ is unchanged by reversing $w$ (the $P_i$ are self-adjoint) and by deleting a letter $A$ at either end ($P_A\xi = \xi$), and $L$ of the empty walk is 1. If $X = P_{u_k} \cdots P_{u_1}\xi$ and $Y = P_{v_l} \cdots P_{v_1}\xi$, so that $u = (u_1, \ldots, u_k)$ and $v = (v_1, \ldots, v_l)$ are their words in chronological order, then

$$\langle X, Y \rangle = L(u_1 \cdots u_k \, v_l \cdots v_1) : \tag{6.2}$$

the word of $X$ followed by the reversed word of $Y$, equal neighbours merged ($P_i^2 = P_i$). A palindromic walk is a word probability: $L(u_1 \cdots u_k \cdots u_1) = \|P_u\xi\|^2 = \Pr_{p_A}(u)$.

**Lemma 6.5** (Relations at short histories)**.** *Assume (H). Let $q$ be one of $A, B, C, D$, let $t$ be a word of "yes" outcomes (in this section a* history *is such a word), and $\eta = P_q P_t \xi$. Then, with the numbers of Definition 5.2,*

*Formalized.* `LengthSeven.WalkForm.shortRules_of_table`; from *e*/3: `LengthSeven.lengthSeven_e3_rules`.

$$(R1_t) \quad P_q P_v P_q \, \eta = o_{qv} \, \eta \qquad (v \neq q, \ |t| \leq 3),$$

$$(R2_t) \quad P_q(P_v P_w + P_w P_v)P_q \, \eta = 2R_{qvw} \, \eta \qquad (v, w, q \text{ distinct}, \ |t| \leq 1),$$

$$(R3_0) \quad P_A(P_x P_y P_z + P_z P_y P_x)P_A \, \xi = 2\omega_y \, \xi \qquad ((x, y, z) \text{ an ordering of } B, C, D),$$

*where $\omega_B = \omega_1$, $\omega_C = \omega_2$, $\omega_D = \omega_3$ (the index is the middle letter). In the probabilities from $p_A$, the words used are $(t, q, v, q)$, $(t, q, v)$, $(t, q)$ for $(R1_t)$; $(t, q, w, q)$, $(t, q, \bar{v}, w, \bar{v}, q)$, $(t, q, v, w, v, q)$, $(t, q, w)$, $(t, q, \bar{v}, w)$, $(t, q, v, w)$, $(t, q)$ for $(R2_t)$; and $(v, \check{v}, A)$, $(v)$ and the empty word for $(R3_0)$, with $v$ one of seventeen words of length at most three. All have length at most six; from e/3 they are the same words after a leading "yes" on $A$, of length at most seven.*

*Proof.* For a word $v$, the purity operator $z_q(v) = P_q E_v P_q - c \, P_q$ of Lemma 4.5 satisfies $z_q(v)^2 = P_q E_v P_q E_v P_q - 2cP_q E_v P_q + c^2 P_q$. Paired with $P_t\xi$ this gives, exactly as in that proof with $\xi$ in place of $\Omega$, $\|z_q(v)P_t\xi\|^2 = \|P_{(t,q,v,\check{v},q)}\xi\|^2 - 2c\|P_{(t,q,v)}\xi\|^2 + c^2\|P_{(t,q)}\xi\|^2$. When the three words have length at most six, (6.1) replaces these by the probabilities from $p_A$, whose combination vanishes, so $z_q(v)\eta = z_q(v)P_t\xi = 0$.

($R1_t$) is this with $v = (v)$ and $c = o_{qv}$; the longest word has length $|t| + 3$.

($R2_t$): in any associative algebra $P_v P_w + P_w P_v = P_w - E_{(\bar{v},w)} + E_{(v,w)}$, where $E_{(\bar{v},w)} = (1 - P_v)P_w(1 - P_v)$ and $E_{(v,w)} = P_v P_w P_v$. Sandwiching by $P_q$ and applying purity to the three terms replaces them by the constants $o_{qw}$, $\langle p_q, U_{e-p_v} p_w \rangle = o_{qw} - 2R_{qvw} + o_{qv}o_{vw}$ (from $U_{e-p} = \mathrm{id} - 2L_p + U_p$) and $\langle p_q, U_{p_v} p_w \rangle = o_{qv}o_{vw}$, whose signed sum is $2R_{qvw}$. The longest words are $(t, q, \bar{v}, w, \bar{v}, q)$ and $(t, q, v, w, v, q)$, of length $|t| + 5$. The letters $v$ and $w$ may be exchanged.

($R3_0$): by the proof of Lemma 5.8, $P_x P_y P_z + P_z P_y P_x$ is a signed sum, with coefficients $\pm\frac{1}{2}$, of the seventeen sequential products $E_v$ with $v \in$

$\{y;\ zy, \bar{z}y;\ xy, \bar{x}y;\ xyz, \bar{x}yz, x\bar{y}z, \bar{x}\bar{y}z;\ zxy, \bar{z}xy, z\bar{x}y, \bar{z}\bar{x}y;\ yxz, \bar{y}xz, y\bar{x}z, \bar{y}\bar{x}z\}$, and the corresponding constants sum to $2Q_{A;xyz} = 2\omega_y$ by (5.1). Here $q = A$, $t$ is empty and $P_A\xi = \xi$; from $p_A$ a leading $A$ costs nothing, so the words $(A, v, \check{v}, A)$, $(A, v)$, $(A)$ are $(v, \check{v}, A)$, $(v)$ and the empty word, of length at most six. □

Since $L$ is unchanged by reversal, each relation may also be applied to the reversed walk; the history is then read from the other end.

**The vectors.** For an ordering $s = (x, y, z)$ of $B, C, D$ and distinct $u, v \in \{B, C, D\}$ put

$$a_x = P_x\xi, \quad \eta^H_s = P_zP_yP_x\xi, \quad \eta^C_s = P_A\eta^H_s, \quad \eta^R_s = P_x\eta^H_s, \quad \eta^J_{uv} = P_vP_u\xi.$$

Lemma C.1 in Appendix C.2 expresses every inner product among these vectors as a polynomial in the thirteen traces. Each entry takes one or two applications of Lemma 6.5; the formulas hold in every representation satisfying (H), for every quadruple, and the octonions enter only through the values of the traces.

**Twenty-eight path vectors.** For a word $w = (w_1, \dots, w_k)$ in the letters $A, B, C, D$ put $[w] = P_{w_k} \cdots P_{w_1}\xi$, and $[\varnothing] = \xi$. The uniform argument uses the set $\mathcal{W}_{28}$ of the twenty-eight words (a list, not $\mathcal{W}_L$ for $L = 28$), written without commas,

$$\begin{array}{ll} \text{ending in } A: & \varnothing,\ CBA,\ DCA,\ BDA,\ DCBA,\ CDBA,\ DBCA; \\ \text{ending in } B: & B,\ CB,\ DB,\ CDAB,\ DCB,\ CDB,\ BDCB; \\ \text{ending in } C: & C,\ DC,\ BC,\ DBAC,\ BDC,\ DBC,\ CBDC; \\ \text{ending in } D: & D,\ BD,\ CD,\ BCAD,\ CBD,\ BCD,\ DCBD. \end{array} \tag{6.3}$$

The words ending in $C$ and $D$ arise from those ending in $B$ by the cyclic relabelling $B \to C \to D \to B$. The $[w]$ are $\xi$, the $a_x$, all six $\eta^J_{uv}$ and $\eta^H_s$, three each of the $\eta^C_s$ and $\eta^R_s$, and the six vectors $[x, z, A]$, $[x, y, A, z]$ with $(x, y, z)$ a cyclic order of $B, C, D$.

**Walk spaces.** The relations of Lemma 6.5 are used below in a second, indefinite space, so the reduction of walks is stated without positivity. A *walk space* is a real vector space $\mathcal{Y}$ with a symmetric bilinear form $b$, linear maps $P_A, \dots, P_D$ with $P_i^2 = P_i$ and $b(P_iu, v) = b(u, P_iv)$, and a vector $\xi = P_A\xi$ with $b(\xi, \xi) = 1$. Walks, histories, $[t] = P_{t_r} \cdots P_{t_1}\xi$ and $L(w) = b(\xi, P_{w_1} \cdots P_{w_n}\xi)$ are defined as for (H). As there, $L$ is unchanged by reversal, by merging equal neighbours and by deleting a letter $A$ at either end, and $b([u], [v]) = L(u\, v^{\mathrm{rev}})$. A walk is *reduced* (for walks) if it has no equal neighbours and neither begins nor ends with $A$.

For a walk $\alpha\, qvq\, \beta$ with $v \neq q$, and for $\alpha\, qvwq\, \beta$ with $q, v, w$ distinct, the relations (R1$_t$) and (R2$_t$) with $t$ the word of $\beta$ read from right to left give

$$L(\alpha\, qvq\, \beta) = o_{qv}\, L(\alpha\, q\, \beta), \qquad L(\alpha\, qvwq\, \beta) = 2R_{qvw}\, L(\alpha\, q\, \beta) - L(\alpha\, qwvq\, \beta). \tag{6.4}$$

The same identities applied to the reversed walk use as history the part $\alpha$, read from left to right; the step is then said to act *from the left*. (R3$_0$) gives, for $\alpha$ empty or ending in $A$,

$$L(\alpha\, xyz) = 2\omega_y\, L(\alpha) - L(\alpha\, zyx). \tag{6.5}$$

**Lemma 6.6** (Reduction of walks)**.** *Fix thirteen real numbers $o_{ij} = o_{ji}$, $R_{ijk}$ ($i, j, k$ distinct letters, $R_{ijk}$ symmetric in its indices) and $\omega_B, \omega_C, \omega_D$. Let a walk space satisfy, for all distinct $x, y, z \in \{B, C, D\}$,* *Not formalized.*

*(a) $L(x) = o_{Ax}$ and $L(xyAzyx) = o_{Ax}o_{xy}R_{Ayz}$;*

*(b) the vector identities ($R1_t$) for $|t| \le 2$, ($R2_t$) for $|t| \le 1$ and ($R3_0$) of Lemma 6.5, for all histories $t$ in the letters $A, B, C, D$, with these thirteen numbers in place of the traces.*

*Then for every walk $w$ of length at most eight there is a polynomial $\Upsilon(w)$ with integer coefficients in the thirteen numbers, depending only on $w$ and not on the walk space, with $L(w) = \Upsilon(w)$. In particular $b([u], [v]) = \Upsilon(u\, v^{\mathrm{rev}})$ for all words $u, v$ of length at most four.*

*Proof.* Each identity below expresses $L(w)$ through 1, the values in (a), the numbers $o_{qv}$, $2R_{qvw}$, $2\omega_y$, and values of walks that are shorter or treated earlier in the list. Induction on the length $n \le 8$ therefore defines $\Upsilon(w)$. It suffices to take $w$ reduced. Then $L(\varnothing) = 1$ and $L(x) = o_{Ax}$; let $n \ge 2$, and let $x, y, z$ be distinct letters of $\{B, C, D\}$ throughout.

*Shortening.* If $w = xAw'$, then ($\mathrm{R1}_0$) at $A$ gives $L(w) = o_{Ax}L(w')$, and likewise at the right end. If $w = \alpha\, qvq\, \beta$, then $|\alpha| + |\beta| = n - 3 \le 5$, so one side has length at most two and the first identity of (6.4), from that side, shortens $w$. From now on $w$ has neither form. Every letter $A$ then stands in places $3, \dots, n-2$, and every factor of $w$ without $A$ is periodic of the form $xyzxyz\cdots$, since its letters differ from both neighbours and from the letters two places away.

*No letter A.* For $n = 2, 3$, ($\mathrm{R2}_0$) at $A$ and ($\mathrm{R3}_0$) paired with $\xi$ give $2L(xy) = 2R_{Axy}$ and $2L(xyz) = 2\omega_y$, using reversal. For $n = 4$, ($\mathrm{R2}_0$) at $x$ gives $2L(xyzx) = 2R_{xyz}o_{Ax}$. For $5 \le n \le 8$, write $w = \alpha\, qvsq$; by periodicity $\alpha$ ends in $s$. The second identity of (6.4) with $\beta$ empty gives $L(w) = 2R_{qvs}L(\alpha q) - L(\alpha\, qsvq)$, and $\alpha\, qsvq$ contains $sqs$.

*One letter A, $n = 5$.* By reversal $w = xyAuv$ with $u \in \{x, z\}$:

$$L(xyAxy) = 2R_{Axy}L(xy) - L(yxAxy), \quad L(xyAxz) = 2R_{Axy}L(xz) - L(xAyxz),$$
$$L(xyAzx) = 2R_{Axz}L(xy) - L(xyAxz), \quad L(xyAzy) = L(yzAyx),$$

by the step at $A$ from the left, at $x$ from the left, at $A$ at the right end, and reversal; the last walk has the second form with the letters renamed.

*One letter A, $n = 6$.* By reversal $w = xyAuvs$ with $\{u, v, s\} = \{x, y, z\}$ and $u \ne y$. The walk $xyAzyx$ is given by (a). For the other three, (6.5) at the right end gives $L(xyAuvs) = 2\omega_v L(xy) - L(xyAsvu)$; the last walk is $xyAzyx$ when $(u, v, s) = (x, y, z)$, and contains $yAy$ otherwise.

*One letter A, $n = 7$, in place four.* Write $w = abcAuvs$ with $\{a, b, c\} = \{u, v, s\} = \{B, C, D\}$ and $u \ne c$. If $s = c$, (6.5) at the right end gives $2\omega_v L(abc) - L(abcAcvu)$, and the last walk contains $cAc$. If $u = a$, the same identity from the left gives $2\omega_b L(avs) - L(cbaAavs)$, which contains $aAa$. The remaining walk is $abcAbca$, and from the left $L(abcAbca) = 2\omega_b L(bca) - L(cbaAbca)$, and $cbaAbca$, whose last letter is its third, is the case $s = c$ above.

*One letter $A$, $n = 7$, in place three.* Now $w = xyAuvsu$ with $u \neq y$. If $u = x$, the step at $x$ from the left gives $2R_{Axy}L(xvsx) - L(xAyxvsx)$. Further

$$L(xyAzyxz) = 2R_{Ayz}L(xyxz) - L(xyzAyxz),$$
$$L(xyAzxyz) = 2R_{xyz}L(xyAz) - L(xyAzyxz),$$

by the step at $y$ from the left with history $x$, and the step at $z$ at the right end; the walk $xyzAyxz$ has its letter $A$ in place four.

*Two letters $A$.* Then $n = 8$ and $w = xyA\,m_1m_2\,Auv$ with $m_1 \neq y$ and $m_2 \neq u$. If $m_1 = x$ or $m_2 = v$, the step at the outer $A$ from that end produces $xAx$ or $vAv$. If $m_2 = y$, the step at $y$ from the left with history $x$ turns the factor $yAm_1y$ into $ym_1Ay$, which produces $AyA$; if $m_1 = u$, the step at $u$ from the right with history $v$ produces $AuA$ in the same way. What remains is $xyAzxAyz$, and the step at $A$ at the right end gives $2R_{Ayz}L(xyAzx) - L(xyAzxAzy)$, a walk with $m_1 = u$.

*One letter $A$, $n = 8$, in place three.* Here $w = xyAuvsuv$, and the step at $v$ at the right end gives $2R_{uvs}L(xyAuv) - L(xyAuvusv)$, which contains $uvu$.

*One letter $A$, $n = 8$, in place four.* Here $w = abcAuvsu$ with $u \neq c$. If $u = a$, (6.5) from the left gives $2\omega_bL(avsa) - L(cbaAavsa)$, which contains $aAa$. Otherwise

$$L(abcAbacb) = 2R_{Abc}L(abacb) - L(abAcbacb),$$
$$L(abcAbcab) = 2R_{abc}L(abcAb) - L(abcAbacb),$$

by the step at $b$ from the left with history $a$, and the step at $b$ at the right end; $abAcbacb$ has its letter $A$ in place three.

These cases exhaust the reduced walks of length at most eight: by reversal a single letter $A$ may be placed in the left half, and two letters $A$ in places $3, \dots, n-2$ at distance at least three occur only in places three and six of a walk of length eight. □

Under (H) a walk space is given by $\mathcal{K}$, and (a), (b) hold: (b) is Lemma 6.5, $L(x) = \Pr_{p_A}(x)$, and $L(xyAzyx)$ is the polarisation $2\langle P_AY, P_AP_zY\rangle = \|P_AY\|^2 + \|P_AP_zY\|^2 - \|P_A(1-P_z)Y\|^2$ with $Y = P_yP_x\xi$, three word probabilities from $p_A$ of length at most four, evaluated with the one-"no" value $\langle p_c, U_{e-p_b}p_a\rangle = o_{ac} - 2R_{abc} + o_{ab}o_{bc}$.

**The comparison form.** Fix four atoms with $\mathcal{A} > 0$. By Theorem 5.5 no two coincide or are orthogonal, so every $o_{ab}$ lies in $(0,1)$. For $x \in \{B, C, D\}$ put

$$m_x = 1/o_{Ax} > 1, \qquad \mu_x = m_x - 1 > 0, \qquad m = m_Bm_Cm_D, \qquad \mu = \mu_B\mu_C\mu_D,$$

and $m_A = 1$. Number the edges of the triangle $BCD$ as $1 = BC$, $2 = CD$, $3 = DB$, so that edge $a$ and the vertex $A$ span the triangle of the holonomy $h_a$ of Section 5.4 with base $j = A$ and order $(k, l, m) = (B, C, D)$; for $x \in \{B, C, D\}$ let $a(x)$ be the edge opposite $x$, so $a(B) = 2$, $a(C) = 3$, $a(D) = 1$. With $\omega_B, \omega_C, \omega_D$ as in Lemma 6.5,

put

$$\hat{c}_1 = R_{ABC} m_B m_C, \quad \hat{c}_2 = R_{ACD} m_C m_D, \quad \hat{c}_3 = R_{ADB} m_D m_B, \quad \hat{c}_4 = R_{BCD}\, m,$$
$$\hat{M}_{11} = o_{BC} m_B m_C - \hat{c}_1^2, \quad \hat{M}_{22} = o_{CD} m_C m_D - \hat{c}_2^2, \quad \hat{M}_{33} = o_{BD} m_D m_B - \hat{c}_3^2,$$
$$\hat{M}_{12} = \hat{c}_1\hat{c}_2 - \omega_C\, m, \quad \hat{M}_{23} = \hat{c}_2\hat{c}_3 - \omega_D\, m, \quad \hat{M}_{13} = \hat{c}_1\hat{c}_3 - \omega_B\, m,$$
$$\hat{S} = \hat{c}_1\hat{c}_2\hat{c}_3 - \hat{c}_1\hat{M}_{23} - \hat{c}_2\hat{M}_{13} - \hat{c}_3\hat{M}_{12}, \qquad \hat{b} = \hat{c}_4 - \hat{S}, \qquad \hat{D} = \det\hat{M} - \hat{b}^2.$$

These are the numbers $c_a$, $M$, $S$, $c_4$ of Section 5.4 in another normalisation: with $f_1 = (o_{BC} m_B m_C)^{1/2}$, $f_2 = (o_{CD} m_C m_D)^{1/2}$, $f_3 = (o_{BD} m_D m_B)^{1/2}$ one has $\hat{c}_a = f_a c_a$ for $a \le 3$, $\hat{M}_{ab} = f_a f_b M_{ab}$ (for instance $f_1 f_2 K_{12} = \omega_C\, m$), $\hat{S} = f_1 f_2 f_3 S$ and $\hat{c}_4 = f_1 f_2 f_3 c_4$. Hence, by Proposition 5.10,

$$\hat{D} = (f_1 f_2 f_3)^2\, \mathcal{D} = o_{BC} o_{CD} o_{BD}\, m^2\, \mathcal{D} = m^3 \mathcal{A} > 0. \tag{6.6}$$

Moreover $\hat{M}$ is positive definite: $M$ is the Gram matrix of the imaginary parts of the octonionic holonomies (end of Section 5.4), so $\hat{M} = \operatorname{diag}(f_a)\, M \operatorname{diag}(f_a)$ is positive semidefinite, and $\det\hat{M} = \hat{D} + \hat{b}^2 > 0$ makes it positive definite.

*The algebra.* Let $C$ be the eight-dimensional real Clifford algebra of the proof of Lemma A.10, generated by $\gamma_1, \gamma_2, \gamma_3$ with $\gamma_k^2 = -1$ and $\gamma_k\gamma_l = -\gamma_l\gamma_k$ $(k \ne l)$. Write $\hat{M} = \Lambda\Lambda^{\mathsf{T}}$ with $\Lambda$ real and invertible, and put $\epsilon_a = \sum_k \Lambda_{ak}\gamma_k$, so that $\epsilon_a\epsilon_b + \epsilon_b\epsilon_a = -2\hat{M}_{ab}$. The elements $\hat{h}_a = \hat{c}_a + \epsilon_a$ $(a = 1, 2, 3)$ play the part of the holonomies. As in the proof of Lemma A.10, $\hat{\Gamma} = \frac{1}{2}(\epsilon_1\epsilon_2\epsilon_3 - \epsilon_3\epsilon_2\epsilon_1) = \det(\Lambda)\gamma_1\gamma_2\gamma_3$ is central, $\hat{\Gamma}^2 = \det\hat{M}$, and $\hat{\Gamma} = \epsilon_1\epsilon_2\epsilon_3 + \hat{M}_{23}\epsilon_1 - \hat{M}_{13}\epsilon_2 + \hat{M}_{12}\epsilon_3$ by (5.12). The eight elements

$$1, \quad \epsilon_1,\ \epsilon_2,\ \epsilon_3, \quad \chi_1 = \epsilon_2\epsilon_3 + \hat{M}_{23},\ \chi_2 = \epsilon_3\epsilon_1 + \hat{M}_{13},\ \chi_3 = \epsilon_1\epsilon_2 + \hat{M}_{12}, \quad \mathfrak{g} = \hat{\Gamma} - \hat{b} \tag{6.7}$$

form a basis, since they differ from the ordered words in the $\epsilon_a$ by a unitriangular change. For $\beta \in C$ write $[\beta]_{\mathfrak{g}}$ and $[\beta]_{\chi_a}$ for its coefficients of $\mathfrak{g}$ and $\chi_a$. Multiplying out with the Clifford relations, $\mathfrak{g}$ occurs in $\epsilon_a$ times a basis element only in

$$\epsilon_a\chi_a = \mathfrak{g} + \hat{b} + \hat{M}_{ac}\epsilon_b - \hat{M}_{ab}\epsilon_c \qquad ((a, b, c) \text{ a cyclic order of } 1, 2, 3), \tag{6.8}$$

so $[\epsilon_a\beta]_{\mathfrak{g}} = [\beta]_{\chi_a}$ for every $\beta$.

*The form on $C$.* Let $u \mapsto \bar{u}$ be the conjugation of $C$, the linear anti-automorphism with $\bar{\gamma}_k = -\gamma_k$; it exists because the defining relations are unchanged by reversing products and changing the signs of the generators. It changes the sign of the $\epsilon_a$ and $\chi_a$ and fixes 1 and $\mathfrak{g}$. Let $\vartheta$ be the linear form with $\vartheta(1) = 1$ that vanishes on the other seven elements of (6.7), and put $\langle u, v\rangle_C = \vartheta(\bar{u}v)$. Since $\vartheta(\bar{u}) = \vartheta(u)$, this form is symmetric and left multiplication by $w$ has adjoint left multiplication by $\bar{w}$. So each $\epsilon_a$ is skew, $\hat{h}_a$ has adjoint $\hat{h}_a^* = \hat{c}_a - \epsilon_a$, and $\hat{h}_a\hat{h}_a^* = \hat{h}_a^*\hat{h}_a = \hat{c}_a^2 + \hat{M}_{aa}$ is a scalar. Only two values of the form are needed: with $\mathfrak{g}_a = -\epsilon_a\mathfrak{g}$,

$$\langle\beta, \mathfrak{g}\rangle_C = \hat{D}\,[\beta]_{\mathfrak{g}}, \qquad \langle\beta, \mathfrak{g}_a\rangle_C = \hat{D}\,[\beta]_{\chi_a} \qquad (\beta \in C). \tag{6.9}$$

Indeed, $\mathfrak{g}$ is central and fixed by conjugation, so for $\beta = \epsilon_a$ or $\chi_a$ the element $\bar{\beta}\mathfrak{g}$ changes sign under conjugation and $\vartheta(\bar{\beta}\mathfrak{g}) = 0$. Also $\vartheta(\mathfrak{g}) = 0$ and, since $\vartheta(\hat{\Gamma}) = \hat{b}$,

$\vartheta(\mathfrak{g}^2) = \det \hat{M} - 2\hat{b}^2 + \hat{b}^2 = \hat{D}$. The second identity follows from the first, since $\langle \beta, \mathfrak{g}_a \rangle_C = \langle \epsilon_a \beta, \mathfrak{g} \rangle_C$.

*The form on four copies of $C$.* Let $\varsigma(\beta) = \beta - 2[\beta]_{\mathfrak{g}}\, \mathfrak{g}$, the reflection of $C$ in the direction $\mathfrak{g}$; by (6.9) it is self-adjoint, and $\varsigma^2 = 1$. On $\mathcal{V} = C^4$, with elements $F = (F_A; F_B; F_C; F_D)$, define the *comparison form*

$$Q(F, F') = \sum_{x,y} \langle F_x, T_{xy} F'_y \rangle_C ,$$

where $T_{AA} = T_{Ax} = T_{xA} = 1$, $T_{xx} = m_x$ for $x \in \{B, C, D\}$, and

$$T_{BC} = \hat{h}_1 \varsigma, \quad T_{CB} = \varsigma \hat{h}_1^*, \quad T_{CD} = \hat{h}_2 \varsigma, \quad T_{DC} = \varsigma \hat{h}_2^*, \quad T_{DB} = \hat{h}_3 \varsigma, \quad T_{BD} = \varsigma \hat{h}_3^*,$$

the elements of $C$ acting by left multiplication. $Q$ is symmetric because $T_{yx} = T_{xy}^*$. For a letter $x$ let $P_x$ be the operator on $\mathcal{V}$ whose $x$-component is $(P_x F)_x = m_x^{-1} \sum_y T_{xy} F_y$ and whose other components vanish. Since $T_{xx} = m_x$, $P_x^2 = P_x$, and $Q(F, P_x F') = m_x^{-1} \sum_{z,y} \langle F_z, T_{zx} T_{xy} F'_y \rangle_C = Q(P_x F, F')$. For a word $w$ put $\hat{u}(w) = P_{w_k} \cdots P_{w_1}(1; 0; 0; 0)$, the counterpart of $[w]$. So $(\mathcal{V}, Q, P_A, \ldots, P_D)$ has the algebraic shape of a representation, but $Q$ is indefinite (Lemma 6.9): it is a bookkeeping device, not a representation. Its blocks are those of an associative model with holonomies $\hat{h}_a$ at $A$, except for the reflection $\varsigma$ (the *twist*) in the direction $\mathfrak{g}$ along the three edges of $BCD$.

**Lemma 6.7** (The comparison form satisfies the relations)**.** *Let $p_A, p_B, p_C, p_D$ be atoms of $J$ with $\mathcal{A} > 0$. Then $(\mathcal{V}, Q, P_A, \ldots, P_D, \hat{u}(\varnothing))$ is a walk space satisfying (a) and (b) of Lemma 6.6 with the thirteen traces of the quadruple. Hence $Q(\hat{u}(u), \hat{u}(v)) = \Upsilon(u\, v^{\mathrm{rev}})$ for all words $u, v$ of length at most four.*

*Narrower in Lean.* `LengthSeven.TableModel.tableModel`, `LengthSeven.TableModel.tableModel_lemmaB`. Lean proves the last sentence for $u, v \in \mathcal{W}_{28}$, entry by entry and for all traces with $o_{AB} o_{AC} o_{AD} \neq 0$, with Lean's Gram table in place of $\Upsilon$; the walk-space assertion is not stated.

*Proof.* The walk-space axioms were noted after the definition of $Q$, and $Q(\hat{u}(\varnothing), \hat{u}(\varnothing)) = \vartheta(1) = 1$. For $F$ with only an $x$-component, $P_y F$ has only the $y$-component $m_y^{-1} T_{yx} F_x$. Two facts are used throughout: $\varsigma$ fixes every element with no $\mathfrak{g}$-coefficient, and an element of $C$ that is a product of at most two of the $\hat{h}_a, \hat{h}_a^*$ has no $\mathfrak{g}$-coefficient, by (6.8). In particular every vector $P_q \hat{u}(t)$ with $|t| \leq 1$ has a $q$-component that is a multiple of 1 or of $T_{qr} 1 \in \{\hat{h}_a, \hat{h}_a^*\}$ ($r \neq q$ in $\{B, C, D\}$), and so lies in the span of 1 and the $\epsilon_a$.

*(R1).* $P_q P_v P_q = m_q^{-1} m_v^{-1} T_{qv} T_{vq} P_q$. If $A \in \{q, v\}$ this is $o_{Ax} P_q$. On the edge $a$ joining $q$ and $v$, $T_{qv} T_{vq}$ is $\hat{h}_a \varsigma^2 \hat{h}_a^*$ or $\varsigma \hat{h}_a^* \hat{h}_a \varsigma$, both equal to $\hat{c}_a^2 + \hat{M}_{aa} = o_{qv} m_q m_v$. So (R1) holds on all of $\mathcal{V}$.

*(a).* $L(x) = m_x^{-1} \vartheta(1) = o_{Ax}$. Let $y, z$ span the edge $a$ and let $\beta$ be the $y$-component of $\hat{u}(xy)$, a multiple $(m_x m_y)^{-1} T_{yx} 1$ of $\hat{h}_c$ or $\hat{h}_c^*$, $c$ the edge $xy$. Then $L(xyAzyx) = Q(\hat{u}(xy), P_A P_z \hat{u}(xy)) = m_z^{-1} \langle \beta, T_{yz} \beta \rangle_C$ (using $T_{zy} = T_{yz}^*$), and since $\varsigma\beta = \beta$ and $\varsigma$ is self-adjoint, $\langle \beta, T_{yz} \beta \rangle_C$ equals $\langle \beta, \hat{h}_a \beta \rangle_C$ or $\langle \beta, \hat{h}_a^* \beta \rangle_C$, which is $\hat{c}_a \langle \beta, \beta \rangle_C$ because $\epsilon_a$ is skew. With $\langle \beta, \beta \rangle_C = o_{xy}/(m_x m_y)$ and $\hat{c}_a = R_{Ayz} m_y m_z$ this is $o_{Ax} o_{xy} R_{Ayz}$.

*(R3$_0$).* The vector $P_A(P_x P_y P_z + P_z P_y P_x)\hat{u}(\varnothing)$ has only the $A$-component $m^{-1}(T_{xy} T_{yz} + T_{zy} T_{yx})1$. Every $\varsigma$ here acts on a product of at most two factors and

can be deleted, which leaves $m^{-1}(\hat{h}_a\hat{h}_b + \hat{h}_b^*\hat{h}_a^*)$ or $m^{-1}(\hat{h}_a^*\hat{h}_b^* + \hat{h}_b\hat{h}_a)$ for the two edges $a, b$ at $y$. By the Clifford relations both equal $2m^{-1}(\hat{c}_a\hat{c}_b - \hat{M}_{ab}) = 2\omega_y$.

*(R2$_t$), $|t| \le 1$.* Let $\beta$ be the $q$-component of $\eta = P_q\hat{u}(t)$. The relation reads $(T_{qv}T_{vw}T_{wq} + T_{qw}T_{wv}T_{vq})\beta = 2R_{qvw}m_qm_vm_w\beta$, with $m_A = 1$. If $q = A$ then $\beta$ is a real multiple of 1 (1, or $m_x^{-1} = o_{Ax}$ for the history $x$), and the relation is linear in $\beta$; on 1 the left side is $(T_{vw} + T_{wv})1 = \hat{h}_a + \hat{h}_a^* = 2\hat{c}_a$. If $v = A$ (or $w = A$) the left side is $(T_{wq} + T_{qw})\beta$; since $\hat{h}_a^*\beta$ is a product of two factors, $\varsigma$ drops out and it equals $(\hat{h}_a + \hat{h}_a^*)\beta = 2\hat{c}_a\beta$. In both cases $\hat{c}_a$ is the required constant.

There remains $\{q, v, w\} = \{B, C, D\}$; by the cyclic relabelling take $q = B$. Write $d_B = \hat{h}_1\hat{h}_2\hat{h}_3$ and cyclically $d_C$, $d_D$, and $n_a = \hat{h}_a\hat{h}_a^*$. The Hermitian part of each $d_x$ is $\hat{c}_4 + \mathfrak{g}$, by the expansion after Theorem 5.9, so $[d_x]_{\mathfrak{g}} = 1$ and $\varsigma(\bar{d}_x) = 2\hat{c}_4 - d_x$. The left side is $\hat{h}_1\varsigma\hat{h}_2\varsigma\hat{h}_3\varsigma\beta + \varsigma\hat{h}_3^*\varsigma\hat{h}_2^*\varsigma\hat{h}_1^*\beta$, and $\beta$ is a multiple of 1, $\hat{h}_3^*$ or $\hat{h}_1$ (histories $\varnothing$, $A$ or $B$; $D$; $C$):

$$\begin{aligned}
\beta = 1: &\quad d_B + \varsigma(\bar{d}_B) = 2\hat{c}_4,\\
\beta = \hat{h}_3^*: &\quad n_3\hat{h}_1\hat{h}_2 + \varsigma\big(\hat{h}_3^*\varsigma(\bar{d}_D)\big) = n_3\hat{h}_1\hat{h}_2 + 2\hat{c}_4\hat{h}_3^* - n_3\hat{h}_1\hat{h}_2 = 2\hat{c}_4\hat{h}_3^*,\\
\beta = \hat{h}_1: &\quad \hat{h}_1\varsigma(d_C) + n_1\hat{h}_3^*\hat{h}_2^* = 2\hat{c}_4\hat{h}_1 - n_1\hat{h}_3^*\hat{h}_2^* + n_1\hat{h}_3^*\hat{h}_2^* = 2\hat{c}_4\hat{h}_1.
\end{aligned}$$

In the second line the outer $\varsigma$ acts on $2\hat{c}_4\hat{h}_3^* - n_3\hat{h}_1\hat{h}_2$, which has no $\mathfrak{g}$-coefficient; in the third, $\varsigma(d_C) = d_C - 2\mathfrak{g} = 2\hat{c}_4 - \bar{d}_C$ and $\hat{h}_1\bar{d}_C = n_1\hat{h}_3^*\hat{h}_2^*$. Since $2\hat{c}_4 = 2R_{BCD}m$, (R2) holds. The last assertion follows from Lemma 6.6 and $\boldsymbol{Q}(\hat{u}(u), \hat{u}(v)) = L(u\,v^{\mathrm{rev}})$. □

**Lemma 6.8** (The words fix the Gram matrix)**.** *Let $p_A, p_B, p_C, p_D$ be atoms of $J$ with $\mathcal{A} > 0$. Let $\mathcal{K}$ be a real Hilbert space with orthogonal projections $P_A, \dots, P_D$ and a unit vector $\xi = P_A\xi$, and let $\mathcal{S}$ be a set of words with $\|P_u\xi\|^2 = \mathrm{Pr}_{p_A}(u)$ for every $u \in \mathcal{S}$. Suppose that*

*(i) $\mathcal{S}$ contains every word of length at most five;*

*(ii) for every history $t$ of length one and distinct $q, v, w$, $\mathcal{S}$ contains the seven words of the instance (R2$_t$) of Lemma 6.5 for at least one choice of the letter that appears twice in its longest word;*

*(iii) $\eta^C_{xyz} + \eta^C_{zyx} = 2\omega_y\,\xi$ for every ordering $(x, y, z)$ of $B, C, D$.*

*Then $\langle[u], [v]\rangle = \boldsymbol{Q}(\hat{u}(u), \hat{u}(v))$ for all words $u, v$ of length at most four, in particular for all $u, v \in \mathcal{W}_{28}$. Hypothesis (H) implies (i)–(iii), by Lemma 6.5.*

*Narrower in Lean.* `LengthSeven.LemmaA.lemmaA_of_gram28Hyp`, `LengthSeven.LemmaA.lemmaA`, with `LengthSeven.TableModel.tableModel`. Lean's conclusion is for $u, v \in \mathcal{W}_{28}$ only, proved entry by entry rather than through Lemma 6.6.

*Proof.* The proof of Lemma 6.5 applies to each relation instance whose words lie in $\mathcal{S}$: by (i) this covers (R1$_t$) with $|t| \le 2$ and (R2$_0$), by (ii) the instances of (R2$_t$) with $|t| = 1$, and (iii) is (R3$_0$). The words behind (a) of Lemma 6.6 have length at most four. So $\mathcal{K}$ satisfies the hypotheses of Lemma 6.6, and $\langle[u], [v]\rangle = \Upsilon(u\,v^{\mathrm{rev}}) = \boldsymbol{Q}(\hat{u}(u), \hat{u}(v))$ by Lemma 6.7. □

Lemma C.1 is not used in the proof of Theorem 6.1. Its derivation uses only words of length at most five, the instances (R2$_t$) with $|t| = 1$ and (R3$_0$), so its formulas hold under the hypotheses (i)–(iii) of Lemma 6.8.

**Lemma 6.9** (Three vectors with negative Gram determinant)**.** *Let $p_A, p_B, p_C, p_D$ be atoms of $J$ with $\mathcal{A} > 0$, and put $\alpha_x = \mu/\mu_x$. There are vectors $\Xi_B, \Xi_C, \Xi_D$ in the real span of the $\hat{u}(w)$, $w \in \mathcal{W}_{28}$, with*

$$Q(\Xi_x, \Xi_x) = 4\hat{D}\alpha_x(\alpha_x + \mu_x), \qquad Q(\Xi_x, \Xi_y) = -4\hat{D}\mu\, m_z \quad (\{x,y,z\} = \{B,C,D\}).$$

*The Gram matrix of the $\Xi_x/(2\alpha_x)$ has determinant $-4\hat{D}^3/\prod_x(1 - o_{Ax}) < 0$.*

*Different route.* `LemmaB.gram_Xi`, `LemmaB.Xi_mem_span_uhat`, `LemmaB.no_psd_realisation_traces`: the Gram entries are checked as a ring identity modulo the trace relation.

*Proof.* Write $\hat{h}'_a = \hat{h}_a - 1$ and $d_a = \hat{h}'_a\hat{h}'^*_a = (\hat{c}_a - 1)^2 + \hat{M}_{aa}$, a scalar. For $F \in \mathcal{V}$ put $(TF)_x = \sum_y T_{xy}F_y$, so that $Q(F', F) = \sum_x \langle F'_x, (TF)_x\rangle_C$; call $TF$ the *source* of $F$.

*Sources.* For $u = (u_B, u_C, u_D) \in C^3$ let $\langle u\rangle = (-u_B - u_C - u_D; u_B; u_C; u_D)$. Then $(T\langle u\rangle)_A = 0$ and $(T\langle u\rangle)_x = (\Theta u)_x$, where $\Theta$ is the $3\times 3$ matrix with entries $\Theta_{xx} = \mu_x$ and $\Theta_{xy} = T_{xy} - 1$. Let $\Theta^0$ be $\Theta$ with $\varsigma$ replaced by 1: its off-diagonal entries are $\hat{h}'_1$ $(BC)$, $\hat{h}'^*_1$ $(CB)$, $\hat{h}'_2$, $\hat{h}'^*_2$, $\hat{h}'_3$ $(DB)$, $\hat{h}'^*_3$ $(BD)$. Since $\varsigma = 1 - 2[\,\cdot\,]_{\mathfrak{g}}\,\mathfrak{g}$ and $[\hat{h}^*_a\beta]_{\mathfrak{g}} = \hat{c}_a[\beta]_{\mathfrak{g}} - [\beta]_{\chi_a}$ by (6.8), the twist changes $\Theta^0$ only through $\mathfrak{g}$-coefficients:

$$(\Theta u)_B = (\Theta^0 u)_B - 2[u_C]_{\mathfrak{g}}\,\hat{h}_1\mathfrak{g} - 2\big(\hat{c}_3[u_D]_{\mathfrak{g}} - [u_D]_{\chi_3}\big)\mathfrak{g}, \tag{6.10}$$

and cyclically for the components $C$ and $D$.

*Where the trace relation enters.* The vector $K_B[\beta]$ below is the column $B$ of the adjugate of $\Theta^0$, with its products ordered so that the $C$- and $D$-components of $\Theta^0 K_B[\beta]$ vanish identically; its $B$-component is the cubic determinant of $\Theta^0$, which the trace relation reduces to $2\mathfrak{g}$. For $\beta \in C$ let $K_B[\beta] \in C^3$ have components

$$u_B = (\mu_C\mu_D - d_2)\beta, \quad u_C = (\hat{h}'_2\hat{h}'_3 - \mu_D\hat{h}'^*_1)\beta, \quad u_D = (\hat{h}'^*_2\hat{h}'^*_1 - \mu_C\hat{h}'_3)\beta,$$

and define $K_C$, $K_D$ by the cyclic relabelling. Because $d_a$ is a scalar, $(\Theta^0 K_B[\beta])_C = (\mu_C\mu_D - d_2 - \mu_C\mu_D + d_2)\hat{h}'^*_1\beta = 0$, and likewise $(\Theta^0 K_B[\beta])_D = 0$. The $B$-component is

$$\big(\mu - \mu_D d_1 - \mu_B d_2 - \mu_C d_3 + \hat{h}'_1\hat{h}'_2\hat{h}'_3 + \hat{h}'^*_3\hat{h}'^*_2\hat{h}'^*_1\big)\beta = (\tau + 2\mathfrak{g})\beta,$$

because the two triple products add up to $2\hat{S}' + 2\hat{\Gamma}$ (in the expansion the terms linear in the $\epsilon_a$ cancel, the quadratic ones give $-2(\hat{c}_a - 1)\hat{M}_{bc}$, the cubic ones $2\hat{\Gamma}$), where $\hat{S}'$ is $\hat{S}$ with every $\hat{c}_a$ replaced by $\hat{c}_a - 1$, and $\tau = \mu - \mu_D d_1 - \mu_B d_2 - \mu_C d_3 + 2(\hat{b} + \hat{S}')$. This is the one place where the trace relation enters: $\tau = 0$. Indeed $\hat{S}' - \hat{S} = \hat{c}_1 + \hat{c}_2 + \hat{c}_3 - 1 - m(\omega_B + \omega_C + \omega_D)$ and $d_1 = o_{BC}m_Bm_C - 2\hat{c}_1 + 1$ (cyclically), so

$$\begin{aligned}\tau = {} & m - \textstyle\sum_{x<y} m_xm_y - m(o_{BC} + o_{CD} + o_{BD}) \\ & + (o_{BC}m_Bm_C + o_{CD}m_Cm_D + o_{BD}m_Bm_D) + 2m\textstyle\sum R - 2m(t_1 + t_2 + t_3),\end{aligned}$$

the sum over the four shape invariants, using $\omega_B + \omega_C + \omega_D = t_1 + t_2 + t_3$. Divided by $m$ (so that $m_xm_y/m = o_{Az}$) this is twice (right side minus left side) of (5.3), which holds for atoms of $J$.

*Six vectors.* Put $k_B = K_B[\epsilon_1]$, $k'_B = K_B[\epsilon_2]$ and, cyclically, $k_C = K_C[\epsilon_2]$, $k'_C = K_C[\epsilon_3]$, $k_D = K_D[\epsilon_3]$, $k'_D = K_D[\epsilon_1]$. Their components are products of at most three of the $\hat{h}'_a$, $\hat{h}'^*_a$ with $\epsilon_1$ or $\epsilon_2$; a $\mathfrak{g}$ appears only through (6.8), from the product $\epsilon_2\epsilon_3\epsilon_1$ in $u_C$ of $k_B$. Expanding in the basis (6.7), the coefficients that enter (6.10) are: for

$k_B$, $[u_C]_{\mathfrak{g}} = 1$, $[u_C]_{\chi_2} = \hat{c}_2 - 1$, $[u_D]_{\chi_3} = \hat{c}_1 - 1$, and no other $\mathfrak{g}$-coefficient; for $k'_B$, no $\mathfrak{g}$-coefficient, $[u_C]_{\chi_2} = 0$ and $[u_D]_{\chi_3} = 1 - \hat{c}_2$; in both, $u_B$ is a multiple of $\epsilon_1$ or $\epsilon_2$. Since $(\Theta^0\langle k_B\rangle)_B = 2\mathfrak{g}\epsilon_1 = -2\mathfrak{g}_1$ and $\hat{h}_1\mathfrak{g} = \hat{c}_1\mathfrak{g} - \mathfrak{g}_1$, (6.10) gives

$$T\langle k_B\rangle = (0;\, -2\mathfrak{g};\, 0;\, -2\mathfrak{g}), \qquad T\langle k'_B\rangle = (0;\, -2\hat{h}'^*_2\mathfrak{g};\, 0;\, 0),$$

and cyclically: $T\langle k_C\rangle$ is $-2\mathfrak{g}$ at $C$ and $B$, $T\langle k_D\rangle$ is $-2\mathfrak{g}$ at $D$ and $C$, $T\langle k'_C\rangle = -2\hat{h}'^*_3\mathfrak{g}$ at $C$, $T\langle k'_D\rangle = -2\hat{h}'^*_1\mathfrak{g}$ at $D$. Also needed are the *auxiliary data* $([\,\cdot\,]_{\mathfrak{g}}, [\,\cdot\,]_{\chi_{a(w)}})$ of the components at $w = B, C, D$: $k_B$ has $(1, 1 - \hat{c}_3)$ at $C$ and $(0, 0)$ elsewhere; $k'_B$ has $(0, \mu_D)$ at $C$, $(0, \mu_C)$ at $D$, $(0, 0)$ at $B$; and cyclically.

*The three vectors.* Put

$$\Xi_B = 2\alpha_B\big((\mathfrak{g}; 0; 0; 0) + \langle k_D\rangle\big) + \mu_B\langle k'_B\rangle - \mu_C\langle k'_C\rangle - \mu_D\langle k'_D\rangle,$$

and $\Xi_C$, $\Xi_D$ cyclically. Since $T(\mathfrak{g}; 0; 0; 0) = (\mathfrak{g}; \mathfrak{g}; \mathfrak{g}; \mathfrak{g})$,

$$T\Xi_B = \big(2\alpha_B\mathfrak{g};\; 2\alpha_B\mathfrak{g} - 2\mu_B\hat{h}'^*_2\mathfrak{g};\; -2\alpha_B\mathfrak{g} + 2\mu_C\hat{h}'^*_3\mathfrak{g};\; -2\alpha_B\mathfrak{g} + 2\mu_D\hat{h}'^*_1\mathfrak{g}\big).$$

The auxiliary data of $\Xi_B$ are $(2\alpha_B, -2\alpha_B\hat{c}_2)$ at $B$ and $(0, 0)$ at $C$ and $D$, the contributions of the $k'$ cancelling in pairs ($\mu_B\mu_D - \mu_D\mu_B$ at $C$); at $A$ the $\mathfrak{g}$-coefficient is $2\alpha_B - 2\alpha_B = 0$.

*Accessibility.* For a word $w$ ending in $f$, the vector $\hat{u}(w)$ has only an $f$-component, namely $T_{w_k w_{k-1}} \cdots T_{w_1 A} 1/\prod_i m_{w_i}$. On each word of (6.3) every $\varsigma$ acts on a product of at most two of the $\hat{h}_a$, $\hat{h}^*_a$, which has no $\mathfrak{g}$-component, so it can be deleted. The products are $1$, $\hat{h}_1$, $\hat{h}_2$, $\hat{h}_3$, $\hat{h}_1\hat{h}_2$, $\hat{h}^*_3\hat{h}^*_2$, $\hat{h}^*_1\hat{h}^*_3$ for the words ending in $A$, and $1$, $\hat{h}_1$, $\hat{h}^*_3$, $\hat{h}^*_2$, $\hat{h}_1\hat{h}_2$, $\hat{h}^*_3\hat{h}^*_2$, $\hat{h}_1\hat{h}_2\hat{h}_3$ for those ending in $B$ (cyclically for $C$, $D$). Their leading terms are $1$, $\pm\epsilon_a$, $\pm\chi_a$ and, for the last word ending in $B$, $\mathfrak{g}$; so the seven components of the words ending in $A$ form a basis of $\{\beta : [\beta]_{\mathfrak{g}} = 0\}$, and those of the words ending in $x \in \{B, C, D\}$ a basis of $\{\beta : [\beta]_{\chi_{a(x)}} = -\hat{c}_{a(x)}[\beta]_{\mathfrak{g}}\}$. By the auxiliary data every component of $\Xi_x$ lies in the corresponding subspace, so $\Xi_x$ is a real combination of the $\hat{u}(w)$.

*The Gram matrix.* Every component of $T\Xi_x$ at $w$ lies in the span of $\mathfrak{g}$ and $\mathfrak{g}_{a(w)}$, because $\hat{h}'^*_a\mathfrak{g} = (\hat{c}_a - 1)\mathfrak{g} + \mathfrak{g}_a$. By (6.9) only auxiliary data enter $Q(\Xi_y, \Xi_x) = \sum_w \langle(\Xi_y)_w, (T\Xi_x)_w\rangle_C$, and $\Xi_B$ has them only at $B$:

$$Q(\Xi_B, \Xi_B) = \hat{D}\big[4\alpha_B^2 - 2\mu_B\big(2(\hat{c}_2 - 1)\alpha_B - 2\alpha_B\hat{c}_2\big)\big] = 4\hat{D}\alpha_B(\alpha_B + \mu_B),$$
$$Q(\Xi_C, \Xi_B) = \hat{D}\big[-4\alpha_B\alpha_C + 2\mu_C\big(2(\hat{c}_3 - 1)\alpha_C - 2\alpha_C\hat{c}_3\big)\big] = -4\hat{D}\mu\, m_D,$$

since $\alpha_C(\alpha_B + \mu_C) = \mu(\mu_D + 1)$, and cyclically.

*The determinant.* For $E_x = \Xi_x/(2\alpha_x)$ and $s_x = \mu_x/\mu^{1/2}$, so that $s_x s_y = 1/\mu_z$, the Gram matrix is $\hat{D}\,G$ with $G_{xx} = 1 + s_x^2$ and $G_{xy} = -(1 + s_x s_y)$. Let $K$ be $G$ with the signs of its off-diagonal entries reversed. It is the Gram matrix of the three vectors $(1, s_x)$ of the plane, so $\det K = 0$. Reversing the off-diagonal signs of a symmetric $3 \times 3$ matrix changes only the term $2K_{BC}K_{CD}K_{BD}$ of its determinant, so

$$\det G = \det K - 4K_{BC}K_{CD}K_{BD} = -4\prod_{x<y}(1 + s_x s_y).$$

Since $1+s_xs_y = m_z/\mu_z = 1/(1-o_{Az})$, the determinant is $-4\hat{D}^3/\prod_z(1-o_{Az})$, negative by (6.6). □

The negative direction comes from the auxiliary element $\mathfrak{g}$, whose square norm is $\hat{D} = m^3\mathcal{A}$: the sources of the $\Xi_x$ are multiples of $\mathfrak{g}$ and $\mathfrak{g}_{a(w)}$, so their Gram matrix is $\hat{D}$ times a matrix that depends only on $o_{AB}, o_{AC}, o_{AD}$. The picture is three vectors of the plane with positive pairwise inner products; the probabilities keep their squared lengths and reverse all three pairwise signs. This is a frustrated triangle, like three spins each required to disagree with both neighbours: reversing one vector flips the signs of two of the three inner products, so the sign of their product is unchanged by any choice of signs, and no choice turns $G$ back into the Gram matrix $K$ of real vectors.

A negative vector can be written down: the adjugate of $G$ has diagonal entries $(s_y - s_z)^2$ and off-diagonal entries $K_{xz}K_{yz} + K_{xy}K_{zz} > 0$ ($\{x,y,z\} = \{B,C,D\}$), so $u = \mathrm{adj}(G)(1,1,1)^{\mathsf{T}}$ has positive entries, and $Gu = (\det G)(1,1,1)^{\mathsf{T}}$ gives $Q\big(\sum_x u_xE_x, \sum_x u_xE_x\big) = \hat{D}\det G\sum_x u_x < 0$. The trace relation is used once, to make $K_x[\beta]$ an exact null direction of the untwisted matrix $\Theta^0$ up to the central element $\mathfrak{g}$; the twist then turns the vectors $K_x[\epsilon_a]$ into sources concentrated on $\mathfrak{g}$.

*Proof of Theorem 6.1.* (a) Either hypothesis gives (H) on the real cyclic representation (Lemma 4.3), as explained above. By Lemma 6.9 there are reals $c_{x,w}$ with $\Xi_x = \sum_{w\in\mathcal{W}_{28}} c_{x,w}\hat{u}(w)$. Put $e_x = \sum_w c_{x,w}[w]$, a vector of the real Hilbert space $\mathcal{K}$. By Lemma 6.8 and bilinearity, $\langle e_x, e_y\rangle = \sum_{w,w'} c_{x,w}c_{y,w'}\,Q(\hat{u}(w),\hat{u}(w')) = Q(\Xi_x,\Xi_y)$. So the Gram matrix of the real vectors $e_x/(2\alpha_x)$ has negative determinant by Lemma 6.9, which is impossible for a Gram matrix of vectors in a real inner product space. The lower bound $L_{\min} \geq 7$ is Proposition 6.2.

(b) Let $\mathcal{W}$ be the finite set of reduced words of length at most seven beginning with "yes" on $A$. The tables on $\mathcal{W}$ of sequential projective representations are restrictions of tables on the set $\mathcal{W}_L$ of Lemma 6.4 with $L = 7$, which form a compact set (Lemma 6.4 and the remark after it). By (a) this set does not contain the octonionic table $f$ from $e/3$, so the continuous function $g \mapsto \max_{w\in\mathcal{W}} |g(w) - f(w)|$ has a positive minimum $\delta$ on it. □

Lemmas 6.6, 6.7, 6.8 and 6.9 are proved here in full. For Lemma C.1, used only for the explicit vectors of Theorem 6.10, Appendix C.2 gives the method and representative entries. The argument gives no explicit value of $\delta$ for a general quadruple, although the vector $u$ of the remark after Lemma 6.9 already carries the contradiction in a single squared norm. The Gram entries of Lemma 6.8 are reached through purity relations, each read from the squared norm of a purity operator applied to a vector, so an error $\varepsilon$ in the table moves them by an amount bounded only by a multiple of $\sqrt{\varepsilon}$, with a constant that depends on every step of the reduction in Lemma 6.6. An inequality linear in the word probabilities with explicit constants for M and S is in the companion paper (section "Explicit margins for M and S").

**Theorem 6.10** (Words of length seven suffice for M and S)**.** *Let $p_A, p_B, p_C, p_D$ be the atoms of* M *or of* S*. The model class is (I). No sequential projective representation, over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$, of any dimension, with projections of any rank and any state, reproduces from $e/3$ the probability of every word of length at most seven that begins with "yes" on $A$; none reproduces from $p_A$ the probability of every word of length at most six. In particular $L_{\min} \le 7$.*

*Different route.* `BlockB.seven_M_e3_algebraic`, `BlockB.seven_S_e3_algebraic`, `BlockB.seven_M_pA_algebraic`, `BlockB.seven_S_pA_algebraic` and the Hilbert and quaternionic versions: the squared norms of the vectors $\Phi$ are evaluated from the Gram entries of Lemma 6.8, not from Lemma C.1.

The theorem is a case of Theorem 6.1(a). Its proof adds a single explicit vector $\Phi = 4\kappa_C + 10\kappa_R - 9\kappa_H$ (for M), built from signed sums $\kappa_X$ of the vectors $\eta_s^X$, whose squared norm the formulas of Lemma C.1 make equal to $-\frac{4}{27}$; for S a similar vector has squared norm $-\frac{2}{729}$. These calculations are in Appendix C.2.

**Proposition 6.11** (Words of length six do not suffice for M and S)**.** *Let $p_A, p_B, p_C, p_D$ be the atoms of* M *or of* S*, and let $\mathsf{Q}_\pm$ be the quaternionic configurations obtained by replacing $v_D$ by $(1,\ j,\ \pm(i+j)/\sqrt{2})$. The equal mixture of the standard representations on $\mathbb{C}^6$ of $\mathsf{Q}_+$ and $\mathsf{Q}_-$, a sequential projective representation on $\mathbb{C}^{12}$ with four projections of rank four and the state* Tr/12 *(Appendix C.5), reproduces the octonionic probability from $e/3$ of every word of length at most six. Hence, with Theorem 6.10, $L_{\min} = 7$ for* M *and for* S*.*

*Narrower in Lean.* `MatchingModels.prop_cer_universalsix_c12`, `MatchingModels.lMin_eq_seven`. Lean gives a model on $\mathbb{C}^{12}$ through length six for every generating quadruple, but not the equal mixture of $\mathsf{Q}_\pm$.

*Proof.* The lower bound can be proved for all words at once. Put $b = \sqrt{2}$ for M, $b = 1$ for S, and $n = 1 + b^2$. Before normalization the columns of the Albert configuration and of its two quaternionic partners are

$$A = (1,0,0), \quad B = (1,b,0), \quad C = (1,i,1), \quad D = (1,j,d),$$

where $d = \ell$ for the Albert configuration and $d = \pm(i+j)/\sqrt{2}$ for the partners. In all three cases the six transition probabilities and four shape invariants are

$$(o_{AB}, o_{AC}, o_{AD}, o_{BC}, o_{BD}, o_{CD}) = (1/n, 1/3, 1/3, 1/3, 1/3, 1/3),$$
$$(R_{ABC}, R_{ABD}, R_{ACD}, R_{BCD}) = (1/(3n), 1/(3n), 1/9, 1/9).$$

For example, the 11 and 22 entries of $p_C \circ p_D$ are $1/9$, and its 12 entry is $[-2(i+j) + d(i-j)]/18$. Its real part vanishes because the coefficients $d_i, d_j$ of $i, j$ in $d$ are equal. Pairing with $p_B$ gives $R_{BCD} = 1/9$.

The double-pair traces are

$$t_1 = \frac{1}{9n}, \qquad t_2 = \frac{1}{9n} + \frac{bd_j}{18n}, \qquad t_3 = \frac{1}{9n} - \frac{bd_i}{18n}.$$

Here is a direct check that also fixes the normalization. The 11,12,13 entries of $p_A \circ p_C$ are $1/3, -i/6, 1/6$; those of $p_B \circ p_D$ are

$$\frac{1}{3n}, \qquad \frac{b}{3n} - \frac{j}{6}, \qquad \frac{-d - bjd}{6n}.$$

The trace pairing gives $t_2$ since $\mathrm{Re}(jd) = -d_j$. Interchanging $C, D$ gives the formula for $t_3$, while $\mathrm{Re}(d(i-j)) = 0$ gives $t_1$. Thus the Albert double-pair vector is the arithmetic mean of the two quaternionic vectors.

The quaternionic partners have the transition probabilities and shape invariants of the Albert configuration, which has $\mathcal{A} > 0$. By the argument of Appendix C.6 (paragraphs "Traces with one occurrence of an atom" and "Expansion and counting"), every word probability of length at most six from $e/3$ is then one and the same affine function of $(t_1, t_2, t_3)$ for the three configurations, so the midpoint identity proves agreement for every word through length six.

The standard embedding $\iota_{\mathbb{H}} : H_3(\mathbb{H}) \to H_6(\mathbb{C})$ preserves quadratic maps and satisfies $\mathrm{Tr} \circ \iota_{\mathbb{H}} = 2\,\mathrm{tr}$. Each quaternionic block is represented by rank-two projections with state $\mathrm{Tr}/6$. Their equal direct sum gives precisely the stated rank-four projections on $\mathbb{C}^{12}$ and state $\mathrm{Tr}/12$. This proves $L_{\min} \geq 7$ for both configurations. The upper bound $L_{\min} \leq 7$ is Theorem 6.10. □

The two blocks have the octonionic transition probabilities and shape invariants, and double-pair traces $\frac{1}{27}(1, \frac{3}{2}, \frac{1}{2})$, $\frac{1}{27}(1, \frac{1}{2}, \frac{3}{2})$ for M and $\frac{1}{18}(1, 1 \pm \frac{\sqrt{2}}{4}, 1 \mp \frac{\sqrt{2}}{4})$ for S, whose averages are the octonionic values.

**Why the mixture fails at seven.** If a word $u = (q_1, \ldots, q_n)$ starts and ends with "yes", its *round trip* is $u^{\mathrm{rt}} = (q_1, \ldots, q_n, q_{n-1}, \ldots, q_1)$, of length $2n - 1$. For atoms in $J$ or $H_3(\mathbb{H})$, purity at the two ends and self-adjointness of the updates give $\Pr_{e/3}(u^{\mathrm{rt}}) = 3\Pr_{e/3}(u)^2$ (the same calculation as Lemma 4.5). Thus, if the two blocks have probabilities $r \pm \varepsilon$ on $u$ and the Albert value is $r$, their equal mixture exceeds the Albert value on $u^{\mathrm{rt}}$ by $3\varepsilon^2$. Take $u = (A, \bar{C}, \bar{B}, D)$. Expanding $U_{e-p} = \mathrm{id} - 2L_p + U_p$ twice,

$$3\Pr_{e/3}(u) = \langle p_D, U_{e-p_B} U_{e-p_C} p_A \rangle = 4\,t_2 + (\text{terms in the } o_{ab} \text{ and } R_{abc}),$$

since the only double-pair trace arises from $4\langle p_D, p_B \circ (p_C \circ p_A)\rangle = 4\langle p_B \circ p_D, p_A \circ p_C\rangle$. The two blocks differ from the octonionic values only in $t_2$ and $t_3$, by $\pm\frac{1}{54}$ for M and $\pm\frac{\sqrt{2}}{72}$ for S (the values displayed above), so $\varepsilon = \frac{4}{3} \cdot \frac{1}{54} = \frac{2}{81}$ for M and $\varepsilon = \frac{4}{3} \cdot \frac{\sqrt{2}}{72} = \frac{\sqrt{2}}{54}$ for S. The excesses on $(A, \bar{C}, \bar{B}, D, \bar{B}, \bar{C}, A)$ are therefore $3\varepsilon^2 = \frac{4}{2187}$ and $\frac{1}{486}$. This explains the failure of this particular mixture; the exclusion of every representation is Theorem 6.1, and the vector of Theorem 6.10 sees the same mechanism at the level of Gram matrices.

*Remark* 6.12 (The obstruction is a variance). The formulas of Lemma C.1 are affine in $\omega_1, \omega_2, \omega_3$ except for the products of two $\omega$'s in three CC entries, which come from the squares of two walk values $L(xyz) = \omega_y$. Summing over the rows, the part of $\|\kappa_C\|^2$ quadratic in the $\omega$'s is

$$\sum_s \left(-2\omega_y^2 + 2\omega_y\omega_z + 2\omega_x\omega_y\right) = 4s_\omega^2 - 8(\omega_1^2 + \omega_2^2 + \omega_3^2), \qquad s_\omega = \omega_1 + \omega_2 + \omega_3,$$

and since $\langle \kappa_H, \kappa_C \rangle = \|\kappa_C\|^2$, $\|\Phi\|^2$ contains it with the factor $16 - 2 \cdot 36 = -56$. The quaternionic partners $\mathsf{Q}_\pm$ of M have the transition probabilities and shape invariants of M and $(\omega_1, \omega_2, \omega_3) = \frac{1}{27}(1, 0, 2)$ or $\frac{1}{27}(1, 2, 0)$, whose mean is the octonionic $\frac{1}{27}(1, 1, 1)$; they do have representations, so the same formulas apply to them. Their affine parts average to the value for M, and exchanging $C$ and $D$

shows that $\mathsf{Q}_+$ and $\mathsf{Q}_-$ give the same value. The quadratic part is $\frac{12}{729}$ for $\mathsf{M}$ and $-\frac{4}{729}$ for $\mathsf{Q}_\pm$, so for the vector $\Phi$ of the proof of Theorem 6.10

$$\|\Phi\|^2(\mathsf{Q}_\pm) = -\frac{4}{27} - 56\Big(-\frac{4}{729} - \frac{12}{729}\Big) = \frac{788}{729} \geq 0.$$

The whole difference between the partners and $\mathsf{M}$ is the term $-8\sum \omega_r^2$, a variance: for fixed $s_\omega$ it is largest when the $\omega_r$ are equal. In the mixture of Proposition 6.11 each block has its own $\omega$'s, and the average of their squares exceeds the square of their average; this is the Gram-level form of the excess $3\varepsilon^2$.

One sequence of seven tests, such as that of the round trip above, never suffices; two exclude the mixtures of quaternionic quadruples. A *schedule* of length seven is a sequence $s = (i_1, \dots, i_7)$ of tests; its words are the $2^7$ assignments of outcomes to $s$. For a pairing $\pi = \{\{i, j\}, \{k, l\}\}$ of $\{A, B, C, D\}$ write $t_\pi = \langle p_i \circ p_j, p_k \circ p_l \rangle$, one of the double-pair traces of Definition 5.2.

**Proposition 6.13** (One schedule does not separate, two defeat quaternionic mixtures)**.** *Let $p_A, p_B, p_C, p_D$ be atoms of $J$ with $\mathcal{A} > 0$; the model class is (I).* *Not formalized.*

*(a) For every schedule $s$ of length seven, a sequential projective representation on $\mathbb{C}^{12}$ reproduces from $e/3$ every word of length at most six and every word of $s$. Consequently a list of words of length at most seven that no sequential projective representation reproduces from $e/3$ contains words of length seven from at least two schedules.*

*(b) Let finitely many quadruples of atoms of $J$, each lying in a unital subalgebra isomorphic to $H_3(\mathbb{H})$ (complex and real quadruples included), be represented by the standard representations on $\mathbb{C}^6$ with the state $\mathrm{Tr}/6$, and mix these blocks with positive weights. The mixture does not reproduce from $e/3$ both every word of length at most five and the two words $(A, \bar{C}, \bar{B}, D, \bar{B}, \bar{C}, A)$ and $(A, \bar{B}, \bar{C}, D, \bar{C}, \bar{B}, A)$.*

*Proof.* (a) Expand a word of $s$ as in the paragraph "Expansion and counting" of Appendix C.6. A term has two factors containing all four atoms only if its split positions strictly between 1 and 7 are position four alone, and only if $i_1, \dots, i_4$ are distinct and $i_4, \dots, i_7$ are distinct. The parts of those two factors that depend on the double-pair traces are then multiples of $t_\pi$ and $t_{\pi'}$, with $\pi = \{\{i_1, i_2\}, \{i_3, i_4\}\}$ and $\pi' = \{\{i_4, i_5\}, \{i_6, i_7\}\}$. Let $\Pi$ be the plane of double-pair traces $t = (t_1, t_2, t_3)$ allowed by the trace relation, $t^J \in \Pi$ the octonionic point (Proposition C.2). If $i_1, \dots, i_4$ or $i_4, \dots, i_7$ repeat a test, let $\Pi_1 \subset \Pi$ be any line through $t^J$; otherwise let $\Pi_1$ be the line through $t^J$ on which $t_\pi$ is constant. On $\Pi_1$ every word probability of $s$ is affine in $t$, as are those of length at most six, so the representation of Proposition C.2(c) for this line reproduces all of them. The consequence holds because a sublist of a list that some representation reproduces is reproduced by it.

(b) Suppose the mixture reproduces these words. In each block $\Pr_{e/3}(u^{\mathrm{rt}}) = 3\Pr_{e/3}(u)^2$ for every word $u$ that begins and ends with "yes" (paragraph "Why the mixture fails at seven"), and so does the octonionic table. The mixture matches $\Pr(u)$ and $\Pr(u^{\mathrm{rt}})$, so the variance of $\Pr(u)$ over the blocks vanishes and every block has the octonionic value of $\Pr(u)$. For $u = (i, j)$ and $u = (i, \bar{j}, k)$, with round trips of

length three and five and $3\Pr_{e/3}(i,\bar{j},k) = o_{ik} - 2R_{ijk} + o_{ij}o_{jk}$, this fixes every $o$ and every $R$ in every block. For $u = (A,\bar{C},\bar{B},D)$ and $u = (A,\bar{B},\bar{C},D)$, $3\Pr_{e/3}(u)$ is $4t_2$, respectively $4t_1$, plus terms in the $o$ and $R$, which fixes $t_2$ and $t_1$, and Proposition 5.3 fixes $t_3$. So each block has the thirteen traces of $p_A,\dots,p_D$, hence $\mathcal{A} > 0$, while its atoms lie in a quaternionic subalgebra, against Theorem 5.5(ii). □

The minimum separating word length depends on the initial state. The analogous minimum from the pure state $p_A$ is six for every quadruple with $\mathcal{A} > 0$: by Theorem 6.1(a) no representation reproduces from $p_A$ every word of length at most six; conversely, conditioning the model of Proposition 6.2 on the outcome "yes" of $A$ reproduces every word of length at most five from $p_A$, because $\Pr_{p_A}(u) = 3\Pr_{e/3}(A,u)$.

*Remark* 6.14 (Faithful and tracial states). For faithful and for corner-tracial states Theorem 6.1(a) has a short direct proof. Let $(\mathcal{K}, P_x, \Omega)$ be the real cyclic representation (Lemma 4.3), $\mathcal{A}_0$ the real $*$-algebra generated by the $P_x$, $M_A = P_A\mathcal{A}_0P_A$ and $\xi = P_A\Omega$. With the notation of Section 5.4 at $j = A$, the elements $H_a - c_aP_A$, $h_a^*h_a - P_A$, $h_ah_a^* - P_A$, $\frac{1}{2}(h_ah_b + h_b^*h_a^*) - K_{ab}P_A$ and $\frac{1}{2}(h_1h_2h_3 + (h_1h_2h_3)^*) - c_4P_A$ of $M_A$ annihilate $\xi$, by Lemma 4.5 applied to words of length at most seven that begin with "yes" on $A$. If the left kernel of $\xi$ in $M_A$ is two-sided, as it is when the state is faithful on $\mathcal{A}_0$ or when $X \mapsto \langle \xi, X\xi\rangle$ is tracial on $M_A$ (for instance for the normalised trace), they vanish on the closure of $M_A\xi$, and the proof of Theorem 5.9 runs there, giving $\mathcal{D} = 0$, against $\mathcal{A} > 0$. The model of Proposition 6.11 is tracial, which is why it must fail at length seven.

### 6.3 Words of length eleven for every configuration

For every configuration with $\mathcal{A} > 0$, the argument of Section 5.4 gives an explicit list of words, the same for every such configuration. A finite list of probabilities cannot give the operator relations (R1)–(R3): by Lemma 4.5, matching the three words $(t,q,v,\check{v},q)$, $(t,q,v)$, $(t,q)$ kills the single vector $\pi(z_q(v))\pi(P_t)\Omega$. Each relation is therefore used only on vectors on which it has been tested. Each of (R1), (R2), (R3) is a signed sum of purity operators $z_q(v)$ with $|v| \le 1, 2, 3$ respectively (proof of Lemma 5.8), so applying it to a vector $\pi(P_t)\Omega$, with $t$ of length $n$ ending in the base of the relation, needs the three words above for each such $z_q(v)$, of lengths up to $n+2$, $n+4$, $n+6$.

Fix the base $j = A$ and the order $(k,l,m) = (B,C,D)$. Applying $h_1 = P_jP_kP_lP_j/\nu_1$ to a vector whose word ends with $j$ appends the letters $l,k,j$, its adjoint appends $k,l,j$, and similarly for $h_2$, $h_3$. A vector $\pi(P_t)\Omega$ has *holonomy degree* $d$ if $t$ is $(j)$ followed by $d$ such blocks; there are $1 + 6 + 36 = 43$ words of degree at most two, of length at most seven. Table 1 lists the relation applications that run the proof of Theorem 5.9 on the single vector $\xi = P_j\Omega$ (notation of Section 5.4). The words these applications require, after merging adjacent equal letters and removing duplicates, form a list $\mathcal{F}_A$: there are 1,673 applications and 2,041 distinct words, and the list is printed in the companion paper (appendix "The eleven-step list").

Table 1: Where the eleven letters go (base $j$). A relation (R1), (R2), (R3) applied to a vector whose word has length $n$ and ends in the base of the relation needs words of length up to $n+2$, $n+4$, $n+6$ (one more if the word does not end in the base).

| Relation needed | Relation used | Vectors | Longest word |
|---|---|---|---|
| $H_a = c_a$ | (R2) at $j$ | holonomy degree $\le 2$ (length $\le 7$) | 11 |
| $h_a^* h_a = h_a h_a^* = 1$ | (R1) $\times 3$ | holonomy degree $\le 2$, and one letter more | 11 |
| Clifford relations (5.10) | (R1), (R3) at $j$ | holonomy degree $\le 1$ (length $\le 4$) | 10 |
| fourth face (5.11) | (R2) at $k$, (R1) | $P_k P_j \Omega$, $P_m P_k P_j \Omega$, $P_l P_k P_j \Omega$, $\xi$ | 6 |
| norm of $\xi$ | none | $P_j \Omega$ | 1 |

**Theorem 6.15** (Words of length eleven suffice)**.** *Let $\mathcal{F}_A$ be the list of* 2,041 *reduced words in four tests $A, B, C, D$ printed in the companion paper (appendix "The eleven-step list"), the words required by the relation applications summarised in Table 1; each begins with "yes" on $A$ and has length at most eleven. Let $p_A, p_B, p_C, p_D$ be atoms of $J$ with $\mathcal{A} > 0$, equivalently (Theorem 5.5) atoms generating $J$; then all six transition probabilities lie in $(0, 1)$. The model class is (I).*

*(a) No sequential projective representation reproduces $\Pr_{e/3}(w)$ for all $w \in \mathcal{F}_A$.*

*(b) Let $\mathcal{F}_B, \mathcal{F}_C, \mathcal{F}_D$ be obtained from $\mathcal{F}_A$ by exchanging the label $A$ with $B$, $C$, $D$, and $\mathcal{F} = \mathcal{F}_A \cup \cdots \cup \mathcal{F}_D$ (8,164 words). For every state $\rho$ of $J$, no sequential projective representation reproduces $\Pr_\rho(w)$ for all $w \in \mathcal{F}$.*

*(c) There are a linear functional $\Lambda$ on the probabilities of the words of $\mathcal{F}$ and the empty word, and $\gamma > 0$, such that $\Lambda \ge 0$ on every sequential projective representation and $\Lambda \le -\gamma$ on the octonionic table from every state.*

*The list does not depend on the atoms. Its words number* 1, 3, 15, 36, 32, 94, 240, 180, 432, 576, 432 *by length* $1, \ldots, 11$.

*Different form in Lean.* `BlockB.thm_eleven_with_overlaps`, which joins `BlockB.thm_eleven_full` and the side assertion `MatchingModels.overlap_pos_lt_one_of_fourPoint_pos`; the counts: `Ends.Eleven.elevenBase_card`, `Ends.Eleven.elevenBase_hist`, `BlockB.elevenUnion_card`. In Lean the list is generated from an encoding of the 1,673 relation applications behind Table 1 (`BlockB.elevenLedgerWords_eq_elevenBase`).

*Proof outline.* The side assertion is Theorem 5.5: no two of the atoms coincide or are orthogonal. (a) Suppose a representation reproduces $\Pr_{e/3}(w)$ for all $w \in \mathcal{F}_A$. By Lemma 4.5 each relation application of Table 1 holds on the vector to which it is applied, since its three words lie in $\mathcal{F}_A$. These applications run the tetrahedron argument (the holonomy form after Theorem 5.9) on the single vector $\xi = P_A \Omega$: the Clifford relations (5.10) and the face relation (5.11) give $\Gamma\xi = (c_4 - S)\xi$, and a Gram computation for $\xi$, $N_a\xi$ and $N_1 N_2 N_3 \xi$ gives $\|\Gamma\xi\|^2 = \det M \, \|\xi\|^2$. Since $\|\xi\|^2 = \frac{1}{3}$, $\mathcal{D} = 0$, against $\mathcal{A} > 0$ and Proposition 5.10. (b) $p_A + p_B + p_C + p_D$ is invertible (proof of Corollary 6.3), so some $\langle \rho, p_j \rangle > 0$, and (a) for the label $j$ applies to the conditioned state as in that proof. (c) follows from (b) as Corollary 4.16 follows from Theorem 4.1, by compactness of the tables of representations on the finite list (Lemma 6.4) and strict separation. The mechanical parts, the list itself, its count and its length histogram, and the verification that each relation application needs only words of the list, are checked in Lean (Appendix G); a proof written for reading is in the companion paper (section "Words of length eleven for every configuration" and appendices "A constructive eleven-step separator" and "The eleven-step list"). □

## 6.4 A positive tetrahedron statistic is not a witness

**Proposition 6.16** ($\mathcal{D} > 0$ is not a witness)**.** *The model class is (I). There is a sequential projective representation on $\mathbb{C}^{12}$, with four projections of rank four and the state* Tr/12 *(Appendix C.5), whose word probabilities agree with those of* S *from e/3 on every word of length at most two, and whose thirteen traces, read from words of length at most four by the formulas of Section 5.2, are those of* S*; so its tetrahedron statistic is $\mathcal{D} = \frac{1}{12} > 0$. The purity operator $T = P_C P_D P_C - \frac{1}{3} P_C$ of the pair $(C, D)$ has*

$$\varphi(T^2) = \tfrac{4}{243},$$

*and the model's deviations from* S *at length three occur exactly on the words whose tests are $C, D, C$ or $D, C, D$, each by $\pm\frac{4}{243}$.*

*Formalized.* `MatchingModels.dfails_length_le_two`, `MatchingModels.dfails_thirteen_traces`, `MatchingModels.dfails_tetra_dtet`, `MatchingModels.dfails_purity_defect`, `MatchingModels.dfails_length_three_deviations`.

*Proof.* Use the matrices $C_1, C_2, C_3$ printed in Appendix C.5 and put $\zeta = C_1C_2C_3 = 1_2 \otimes \sigma_x$. They are skew-Hermitian, square to $-1$ and anticommute. Thus $\zeta$ is a central self-adjoint involution with two-dimensional $+1$ and $-1$ eigenspaces. The normalized trace $\tau = \mathrm{Tr}_4/4$ vanishes on $\zeta$, on each $C_i$ and on each $C_iC_j$ with $i \neq j$. For the latter two assertions, conjugation by a suitable $C_k$ changes the sign; for $\zeta$ use its displayed tensor form.

The four column isometries defining the projections are

$$V_A = (1,0,0)^{\mathsf{T}}, \quad V_B = (1,1,0)^{\mathsf{T}}/\sqrt{2},$$
$$V_C = (1,-C_1,1)^{\mathsf{T}}/\sqrt{3}, \quad V_D = (1,-C_2,-C_3)^{\mathsf{T}}/\sqrt{3}.$$

where each entry is a $4\times4$ block. They satisfy $V_i^*V_i = 1_4$, so $P_i = V_iV_i^*$ are rank-four projections. With the $4\times4$ blocks $K_{ij} = V_i^*V_j$,

$$K_{AB} = 1/\sqrt{2}, \quad K_{AC} = K_{AD} = 1/\sqrt{3}, \quad K_{BC} = (1-C_1)/\sqrt{6}, \quad K_{BD} = (1-C_2)/\sqrt{6},$$
$$K_{CD} = (1 - C_1C_2 - C_3)/3 = (1 + (\zeta - 1)C_3)/3.$$

For all pairs except $CD$, $K_{ij}K_{ji} = o_{ij}1_4$, with $o_{AB} = 1/2$ and $o_{ij} = 1/3$ otherwise. The remaining product is

$$K_{CD}K_{DC} = (3 - 2\zeta)/9.$$

Its scalar values in the two sectors are $1/9$ and $5/9$. Their equal-weight mean is $1/3$ and their variance is $4/81$. Consequently all yes–yes probabilities through length two agree with S; complementation and the common one-test marginals give every other outcome choice.

*The thirteen traces.* Trace cyclicity gives, for any cyclic word,

$$3\,\mathrm{Re}\,\varphi(P_{i_1}\cdots P_{i_n}) = \mathrm{Re}\,\tau(K_{i_1i_2}\cdots K_{i_ni_1}).$$

The signed-update formulas defining the thirteen traces therefore give their usual pair, triangle and symmetrized four-product expressions. The shape invariants involving $A$ are $1/6, 1/6, 1/9$. For $BCD$, the numerator is

$$(1-C_1)(1-C_1C_2-C_3)(1+C_2) = 2 - \zeta - 2C_1C_2 - C_3 + C_1C_3 + C_2C_3.$$

Its normalized real trace is two and the denominator is eighteen, giving $R_{BCD} = 1/9$. The three $\omega$'s are the real cyclic products with label orders $ACBD$, $ABCD$ and $ACDB$. Their numerators, each over eighteen, are

$$(1 + C_1)(1 - C_2), \qquad (1 - C_1)(1 - C_1C_2 - C_3), \qquad (1 - C_1C_2 - C_3)(1 + C_2).$$

Each has normalized real trace one. Thus all $\omega_r = 1/18$, and $2t_1 = \omega_2 + \omega_3$ and its cyclic counterparts give all $t_r = 1/18$. At base $A$ the resulting data have

$$(c_1, c_2, c_3, c_4) = (1/\sqrt{2}, 1/\sqrt{3}, 1/\sqrt{2}, 1/\sqrt{3}), \qquad M = \operatorname{diag}(1/2, 2/3, 1/2).$$

Hence $S = c_1c_2c_3 = 1/(2\sqrt{3})$ and $\mathcal{D} = 1/6 - 1/12 = 1/12$.

*Purity and the three-letter deviations.* The purity operator is

$$T = V_C(-2\zeta/9)V_C^*, \qquad T^2 = (4/81)P_C,$$

so $\varphi(T^2) = 4/243$. In each sector all pairs satisfy $P_iP_jP_i = \lambda_{ij}P_i$. A three-yes word $i, j, k$ therefore has probability $\lambda_{ij}\lambda_{jk}/3$, while the Albert probability is $o_{ij}o_{jk}/3$. Only $\lambda_{CD}$ varies with the sector. The average differs precisely when both edges are $CD$, namely for $CDC$ and $DCD$, and the difference is $\operatorname{Var}(\lambda_{CD})/3 = 4/243$.

This also determines every yes–no choice without listing words. With the tracial initial state, write the probability as $\langle q_i, U_{q_j}q_k\rangle/3$, using the Albert trace pairing on one side and $\langle x, y\rangle = 3\varphi(x \circ y)$ on the matrix side. Complementing either endpoint is linear. Complementing the middle outcome uses $U_{1-p} = I - 2L_p + U_p$; the additional terms depend only on the single-test, pair traces and shape invariants, which have just been shown equal. Thus the discrepancy is the three-yes discrepancy times the signs of the two endpoint outcomes (+1 for yes, −1 for no); the middle outcome does not change it. All eight choices for each of $CDC, DCD$ therefore differ by $\pm 4/243$, and no other pattern differs. Consecutive repeated tests reduce to at most two tests or to zero, so unreduced words cause no exception. □

The model is a mixture of two sectors on which the four projections act as quaternionic configurations whose fourth atoms are rotated in opposite directions; the mixture restores the double-pair traces but cannot keep every conditional state pure. Hence the separating words of Theorem 6.1 include words with three intermediate outcomes.

## 6.5 No linear equality at short length

Satisfying every linear equality only places a table in the affine hull of the model tables (the empty word supplies the constant term); belonging to the convex set of model tables is a stronger property, and it is the one that fails at seven (Theorem 6.1).

The certificates are inequalities. A linear *equality* $\sum_w c_w \Pr(w) = 0$, valid in every sequential projective representation and violated by $J$, would also separate. This is the form an $s$-identity (a Jordan polynomial identity that holds in every special Jordan algebra but not in all Jordan algebras[60]) such as Glennie's identity[61]

[60] K. McCrimmon, The role of identities in Jordan algebras, *Resenhas IME-USP* **6** (2004), no. 2/3, 265–280. doi:10.11606/resimeusp.v6i2/3.75294.

[61] Glennie, "Some identities valid in special Jordan algebras", note 7.

would take if it could be read as a relation among word statistics. Call $c$ a *projective equality* if $\sum_w c_w \varphi(E_w) = 0$ in every sequential projective representation. Let $FJ_n$ be the free unital Jordan algebra on idempotents $p_1, \dots, p_n$, $\omega : FJ_n \to \mathbb{R}[\mathbb{Z}_2^{*n}]^+$ the homomorphism with $p_i \mapsto \frac{1}{2}(1 + r_i)$, and $V_n \subseteq FJ_n$ the span of the elements $U_{q_1} \cdots U_{q_k} 1$, $q_t \in \{p_i, 1-p_i\}$. For a combination $c$ put $f(c) = \sum_w c_w U_{q_1} \cdots U_{q_k} 1 \in V_n$, so that $\omega(f(c)) = E(c)$.

**Proposition 6.17** (Linear equalities)**.** *Fix $n$ tests; the model class is (I).*

*(a) $c$ is a projective equality if and only if $E(c) = 0$ in $\mathbb{R}[\mathbb{Z}_2^{*n}]$; it suffices to test real finite-dimensional representations with vector states.*

*(b) Every $c$ with $f(c) = 0$ in $FJ_n$ is a projective equality and holds for idempotents of every Jordan algebra and every state. These include the arrow-of-time equalities $\Pr(u, i) + \Pr(u, \bar{\imath}) = \Pr(u)$ and the repetition equalities $\Pr(u, a, a, v) = \Pr(u, a, v)$ and $\Pr(u, i, \bar{\imath}, v) = 0$, but are not generated by them: for $i \neq j$,*

$$\Pr(\bar{\imath}, j) - \Pr(i, j) - \Pr(j) = \Pr(\bar{\jmath}, i) - \Pr(j, i) - \Pr(i), \qquad \Pr(i, j, \bar{\imath}) = \Pr(i, \bar{\jmath}, \bar{\imath}). \tag{6.11}$$

*(c) Let $p_1, \dots, p_n$ be idempotents of $J$. A projective equality violated by their table from some state exists if and only if some $x \in V_n$ has $\omega(x) = 0$ and $x(p_1, \dots, p_n) \neq 0$. If the $p_i$ generate $J$, some $x \in FJ_n$ has these properties; so if $V_n = FJ_n$, such an equality exists.*

*(d) For any four atoms of $J$ with $\mathcal{A} > 0$, no projective equality supported on words of length at most six is violated from $e/3$.*

*Narrower in Lean.* (a) `AppF.prop_cer_noequality_a`, with projective equalities defined by real finite-dimensional vector states and shown equivalent to vanishing on algebraic models in Lean's lowest universe, through which representations enter; (d) `MatchingModels.no_classI_linearEq_le_six`. Lean has no (b), (c).

*Proof.* Parts (a) and (b) are proved in Appendix C.3.

(c) The octonionic table is $\Pr_\rho(w) = \langle \rho, f(w)(p) \rangle$, so $c$ is violated for some $\rho$ iff $f(c)(p) \neq 0$, and by (a) it is a projective equality iff $\omega(f(c)) = 0$. If the atoms generate $J$ and every $x \in \ker \omega$ had $x(p) = 0$, evaluation at $p$, which maps $FJ_n$ onto $J$, would factor through the special Jordan algebra $\omega(FJ_n)$; but $J$ is not a homomorphic image of a special Jordan algebra (Albert and Paige[62]; see also Glennie's proof[63]).

(d) For four atoms with $\mathcal{A} > 0$, the probabilities from $e/3$ of the words of length at most six are reproduced by the representation on $\mathbb{C}^{12}$ of Proposition 6.2 (for M and S also by that of Proposition 6.11), and every sequential projective representation satisfies every projective equality. □

[62] A. A. Albert and L. J. Paige, On a homomorphism property of certain Jordan algebras, *Trans. Amer. Math. Soc.* **93** (1959) 20–29. doi:10.1090/S0002-9947-1959-0108524-2.

[63] Glennie, "Some identities valid in special Jordan algebras", note 7. Thm. 4(a), pp. 58–59.

*Remark* 6.18 (Equalities through five steps). For any four atoms of $J$, in particular those of $\mathsf{S}$, no projective equality supported on words of length at most five is violated from any state of $J$. By Proposition 6.17(a) this says that every combination $c$ supported on words of length at most five with $E(c) = 0$ has $\sum_w c_w \Pr_\rho(w) = 0$ for every state $\rho$ of $J$. It is proved in the companion paper (section "No equality through five steps" and appendix "Equalities through five steps"), and it is machine-checked in both forms (Appendix G). Both proofs show $f(c)(p) = 0$, and the idea is as follows. Each word becomes a combination of expressions built from the operators $2L_{r_i}$ and $U_{r_i}$ of the symmetries $r_i = 2p_i - e$ (elements with square $e$). Those of degree nine are single reflections with linearly independent special values, so they drop out. Those of degree eight are products of two reflections, and either cancel class by class or contract, without changing their value in $J$, to lower degree. Below degree eight, a top component involving all four atoms has one atom occurring only once; the other three lie in one complex subalgebra (Lemma A.3), and the associative action of its matrix algebra on the octonionic matrices carries a relation that holds in the group algebra over to $J$. The needed lower-degree corrections are written down as Jordan polynomials.

*Formalized.* `HorizonFive.horizonFive_projectiveEquality`, assuming only that $c$ vanishes on real finite-dimensional representations with vector states; the form with $E(c) = 0$: `HorizonFive.horizonFiveClause`.

For general instruments, with independent settings at each time step, the arrow-of-time equalities are the only linear equalities (Clemente and Kofler[64]; Hoffmann, Spee, Gühne and Budroni[65]); part (b) shows that a fixed projection per test with the Lüders rule forces more.

Whether a separating equality exists at greater length is open; by (c) it is tied to whether the elements $U_{q_1} \cdots U_{q_k} 1$ span the free Jordan algebra on idempotents. An $s$-identity outside $V_n$ has no reading as a combination of word probabilities. Near a complex subalgebra every such equality is weak: its residual vanishes to second order, and states in the subalgebra see it only at third order; this is proved in the companion paper (section "The second-order neighbourhood of a complex subalgebra").

No link is established here between the length seven of Theorem 6.1 and the degree eight of Glennie's identity. They are facts of different kinds about the same quadruples. The obstruction of Theorem 6.1 is one of positivity (a purity operator with three intermediate outcomes needs words of length $1+3+2+1 = 7$), and every such obstruction needs four tests (Proposition 5.1). Theorem F.6 (Appendix F) is of equality type: a Jordan polynomial of degree two in each of four variables that vanishes in every special Jordan algebra is nonzero at the four symmetries exactly when $\mathcal{A} > 0$. Whether that polynomial is the value $f(c)$ of a combination $c$ of word probabilities, and of which length, is not decided here.

## 6.6 Robustness

A numerical optimisation (not a proof) indicates that the margin of Theorem 6.1(b), the least maximal word-probability error on its word list that a representation can

[64] L. Clemente and J. Kofler, No fine theorem for macrorealism: limitations of the Leggett–Garg inequality, *Phys. Rev. Lett.* **116** (2016) 150401. arXiv:1509.00348.

[65] J. Hoffmann, C. Spee, O. Gühne and C. Budroni, Structure of temporal correlations of a qubit, *New J. Phys.* **20** (2018) 102001. arXiv:1712.01234.

attain, is about $6.0\times10^{-5}$ for M and about $5\times10^{-5}$ for S. The proof of Theorem 6.1(b) gives no value. A separating inequality with known coefficients gives a proved lower bound, by the next proposition; with the inequality of the companion paper (section "Explicit margins for M and S") the proved margins are above $5.1\times10^{-12}$ for M and above $7.7\times10^{-13}$ for S.

**Proposition 6.19** (Robust version)**.** *Let $\Lambda=\sum_w \lambda_w[w]$ (a* witness*) be supported on words of length at most $L$, with $\Lambda(g)\ge 0$ for the table $g$ of every sequential projective representation and $\Lambda(f)\le-\gamma<0$ for a target table $f$, and put $\|\lambda\|_1=\sum_w|\lambda_w|$. Every sequential projective representation has $|g(w)-f(w)|\ge\gamma/\|\lambda\|_1$ on some word $w$ of the support; if $|f'(w)-f(w)|\le\delta$ on the support, it misses $f'$ by at least $\gamma/\|\lambda\|_1-\delta$.*

*Narrower in Lean.* `FiniteRealWordFunctional.separation_forces_error`. The first assertion; the second, for a target known within $\delta$, is not stated.

*Proof.* $0\le\Lambda(g)=\Lambda(f)+\Lambda(g-f)\le-\gamma+\|\lambda\|_1\max|g-f|$, and the triangle inequality. □

The next result concerns instruments with arbitrary positive branches, provided their disturbance is uniformly controlled on both spectral subspaces of the effect rounded to the projection $1_{[1/2,1]}(A)$; it does not infer that control from the observed word probabilities.

**Proposition 6.20** (Stability under uniform face calibration)**.** *Let $(\mathcal{E}_1,\mathcal{E}_0)$ be bounded positive real-linear maps on the self-adjoint trace-class operators of a real, complex or quaternionic Hilbert space, with trace-preserving sum. Write $A=\mathcal{E}_1^*(1)$, suppose $\|A-A^2\|\le\varepsilon<\frac14$, and set*

$$\tau=\tfrac12(1-\sqrt{1-4\varepsilon}),\qquad P=1_{[1/2,1]}(A),\qquad Q=1-P.$$

*Suppose, for some $\delta\ge0$, that*

$$\|\mathcal{E}_1(\sigma)-\sigma\|_1\le\delta\quad(\sigma=P\sigma P),\qquad\|\mathcal{E}_0(\sigma)-\sigma\|_1\le\delta\quad(\sigma=Q\sigma Q)\tag{6.12}$$

*for every density operator on the indicated subspace. Put*

$$b=\frac{\delta+\tau+\sqrt{(\delta-\tau)^2+4\tau}}{2},\qquad\beta=\min\{2,b\}.$$

*Then, for every density operator $\rho$,*

$$\|\mathcal{E}_1(\rho)-P\rho P\|_1\le\beta,\qquad\|\mathcal{E}_0(\rho)-Q\rho Q\|_1\le\beta.\tag{6.13}$$

*For a family of fixed instruments satisfying these hypotheses with the same $\delta,\varepsilon$, the word probabilities from any fixed density operator differ from those of the rounded projective model by at most $n\beta$ on a word of length $n$. Consequently, with the witness and target of Proposition 6.19, if the instrument table is within $\eta$ of the target on the witness support, then*

$$\eta+L\beta\ge\gamma/\|\lambda\|_1.\tag{6.14}$$

*Narrower in Lean.* `calibrated_positive_instrument_word_error_le`. The word bound $n\beta$ in an abstract trace-norm setting, with the leakage bound $\tau$ as a hypothesis; the trace-class Hilbert-space instance and (6.14) are not formalized.

*Proof.* The first step controls the off-diagonal part using positivity alone. For a positive map $\mathcal{E}$ and $\rho \geq 0$, put

$$U = \mathcal{E}(P\rho P), \quad V = \mathcal{E}(Q\rho Q), \quad H = \mathcal{E}(P\rho Q + Q\rho P).$$

For every real $t$, positivity gives $t^2U + tH + V = \mathcal{E}((tP + Q)\rho(tP + Q)) \geq 0$. Let $F_+, F_-$ be the positive and negative spectral support projections of $H$. Taking the real trace against either $F = F_\pm$ and using the discriminant of this nonnegative scalar quadratic gives $|\operatorname{Tr}(FH)| \leq 2\sqrt{\operatorname{Tr}(FU)\operatorname{Tr}(FV)}$. This also holds when a quadratic coefficient vanishes. Adding the two inequalities, then applying scalar Cauchy–Schwarz and $F_+ + F_- \leq 1$, yields

$$\|H\|_1 \leq 2\sqrt{\operatorname{Tr} U \ \operatorname{Tr} V}. \tag{6.15}$$

The argument uses only self-adjoint trace-class operators and the real trace, so applies over all three scalar fields and in arbitrary dimension.

Now let $\rho$ have trace one and put $p = \operatorname{Tr}(P\rho P)$, $q = 1 - p$. The spectral bound on $A$ gives $QAQ \leq \tau Q$. For $\mathcal{E} = \mathcal{E}_1$, therefore, $\operatorname{Tr} U \leq p$ and $\operatorname{Tr} V = \operatorname{Tr}(AQ\rho Q) \leq \tau q$. By (6.12), homogeneity and (6.15),

$$\|\mathcal{E}_1(\rho) - P\rho P\|_1 \leq \delta p + \tau q + 2\sqrt{\tau p q} \leq b.$$

The last bound is the largest eigenvalue of $\left(\begin{smallmatrix} \delta & \sqrt{\tau} \\ \sqrt{\tau} & \tau \end{smallmatrix}\right)$, whose quadratic form is evaluated at $(\sqrt{p}, \sqrt{q})$. Interchanging $P, Q$ and $A, 1 - A$ gives the other branch. Both outputs being positive with trace at most one, their trace-norm distance is also at most two; this gives $\beta$.

Every positive trace-nonincreasing map contracts trace norm on self-adjoint trace-class operators: apply it to the positive and negative parts and use the triangle inequality. By homogeneity, (6.13) has upper bound $\beta \operatorname{Tr} X$ on any $X \geq 0$. Telescoping the two compositions along a word, with the same initial density operator, thus bounds their output trace-norm difference by $n\beta$. Taking traces gives the word-probability bound, and Proposition 6.19 gives (6.14). □

The calibration in (6.12) concerns every state on the spectral subspaces of the rounded effect, not merely the exact certainty faces of $A$, a finite set of preparations, or states reached in one experiment. The conclusion is a dimension-independent bound for unnormalised branches and word probabilities from density operators; no bound on normalised postselection or on ancillary extensions is asserted. At $\delta = \varepsilon = 0$ it recovers the Lüders rule. For small errors $b \leq \delta + \tau + \sqrt{\tau}$, so the estimate tends to zero.

The word-probability bound also combines directly with Theorem 6.1(b), which needs no explicit inequality. Let the atoms satisfy $\mathcal{A} > 0$, let $\delta_0 > 0$ be the margin of Theorem 6.1(b), and suppose that instruments as in the proposition, from a density operator $\rho$, give a table within $\eta$ of the octonionic table from $e/3$ on the words of length at most seven that begin with "yes" on $A$. The rounded projections with the state $\rho$ form a sequential projective representation, which differs from the

octonionic table by at least $\delta_0$ on one of these words and from the instrument table by at most $7\beta$ on each. Hence

$$\eta + 7\beta \geq \delta_0. \tag{6.16}$$

When the effects are exact projections, a match on the words of length at most seven forces some test, on a state where its answer is certain, to change the probability of a later word.

*Not formalized.*

**Proposition 6.21** (Sharp effects disturb a certain answer)**.** *Let $p_A, p_B, p_C, p_D$ be atoms of $J$ with $\mathcal{A} > 0$, and let $\delta_0 > 0$ be such that every sequential projective representation differs from the octonionic table from $e/3$ by at least $\delta_0$ on some word of length at most seven beginning with "yes" on $A$ (Theorem 6.1(b)). On a real, complex or quaternionic Hilbert space of any dimension let $\rho_0$ be a density operator and, for each test $i$, let $(\mathcal{E}_{i,1}, \mathcal{E}_{i,0})$ be bounded positive real-linear maps on the self-adjoint trace-class operators, with trace-preserving sum and with $\mathcal{E}_{i,1}^*(1) = P_i$ a projection. Write $\Pr_\sigma(q_1, \dots, q_n) = \operatorname{Tr}(\mathcal{E}_{q_n} \cdots \mathcal{E}_{q_1} \sigma)$. If $\Pr_{\rho_0}$ equals the octonionic table on all words of length at most seven beginning with "yes" on $A$, then there are an outcome $q$ of some test, a density operator $\sigma$ with $\Pr_\sigma(q) = 1$, and a nonempty word $u$ of length at most six with*

$$|\Pr_\sigma(q, u) - \Pr_\sigma(u)| \geq \delta_0/2; \qquad \textit{in particular} \quad \|\mathcal{E}_q(\sigma) - \sigma\|_1 \geq \delta_0.$$

*Proof.* Put $\mathcal{L}_{i,1}X = P_i X P_i$ and $\mathcal{L}_{i,0}X = (1 - P_i)X(1 - P_i)$. For $X \geq 0$, $\mathcal{E}_{i,1}((1 - P_i)X(1 - P_i)) \geq 0$ has trace $\operatorname{Tr}(P_i(1 - P_i)X(1 - P_i)) = 0$, so it vanishes, and (6.15) then kills $\mathcal{E}_{i,1}(P_i X(1 - P_i) + (1 - P_i)XP_i)$. Hence $\mathcal{E}_q = \mathcal{E}_q \mathcal{L}_q$ for every outcome $q$ (for "no" by the same argument with $1 - P_i = \mathcal{E}_{i,0}^*(1)$), and $\operatorname{Tr} \mathcal{E}_q = \operatorname{Tr} \mathcal{L}_q$.

The projections $P_i$ with the state $\operatorname{Tr}(\rho_0 \cdot)$ form a sequential projective representation, so some word $w = (q_1, \dots, q_n)$ of the stated kind has $|\Pr_{\rho_0}(w) - \operatorname{Tr}(\mathcal{L}_{q_n} \cdots \mathcal{L}_{q_1} \rho_0)| \geq \delta_0$. Let $Y_k = \mathcal{L}_{q_k} \cdots \mathcal{L}_{q_1} \rho_0$, $\Phi_k = \mathcal{E}_{q_n} \cdots \mathcal{E}_{q_{k+1}}$ and $h_k = \operatorname{Tr} \Phi_k Y_k$. Then $h_0 = \Pr_{\rho_0}(w)$, $h_n$ is the projective value, and $h_{n-1} = h_n$ because $\operatorname{Tr} \mathcal{E}_{q_n} = \operatorname{Tr} \mathcal{L}_{q_n}$. Since $\mathcal{E}_{q_k} Y_{k-1} = \mathcal{E}_{q_k} Y_k$,

$$h_{k-1} - h_k = \operatorname{Tr} \Phi_k \big(\mathcal{E}_{q_k} Y_k - Y_k\big), \qquad \operatorname{Tr} Y_k \leq \operatorname{Tr}(P_A \rho_0) = \Pr_{\rho_0}(A) = \tfrac{1}{3}.$$

The at most six differences with $1 \leq k \leq n - 1$ sum to at least $\delta_0$ in absolute value, so one of them is at least $\delta_0/6$ in absolute value. Then $\operatorname{Tr} Y_k > 0$. Put $\sigma = Y_k / \operatorname{Tr} Y_k$, $q = q_k$ and $u = (q_{k+1}, \dots, q_n)$, a nonempty word since $k \leq n - 1$. As $\operatorname{Tr} \mathcal{E}_{q_k} = \operatorname{Tr} \mathcal{L}_{q_k}$ and $\mathcal{L}_{q_k} Y_k = Y_k$, $\Pr_\sigma(q) = \operatorname{Tr} \mathcal{L}_{q_k} \sigma = 1$, and

$$\Pr_\sigma(q, u) - \Pr_\sigma(u) = \frac{\operatorname{Tr} \Phi_k(\mathcal{E}_{q_k} Y_k - Y_k)}{\operatorname{Tr} Y_k} = \frac{h_{k-1} - h_k}{\operatorname{Tr} Y_k}.$$

Dividing a difference of absolute value at least $\delta_0/6$ by $\operatorname{Tr} Y_k \leq \frac{1}{3}$ gives $|\Pr_\sigma(q, u) - \Pr_\sigma(u)| \geq \delta_0/2$, the first bound. The operator $\mathcal{E}_q(\sigma) - \sigma$ has trace zero and $0 \leq \Phi_k^*(1) \leq 1$, so $|\operatorname{Tr} \Phi_k(\mathcal{E}_q(\sigma) - \sigma)| \leq \frac{1}{2}\|\mathcal{E}_q(\sigma) - \sigma\|_1$, which gives the second. □

So, in the notation of Proposition 6.20 with $\varepsilon = 0$, such instruments have calibration constant $\delta \geq \delta_0$, whereas (6.16) alone gives only $\delta \geq \delta_0/7$. Instruments of this kind that reproduce the octonionic table on every word exist (Corollary 8.7; Example 8.9 shows one disturbed certain answer).

*Remark* 6.22 (The disturbed state is not prepared by the instruments). Suppose the instruments of Proposition 6.21 reproduce the octonionic table from $e/3$ on every finite word. Let $K$ be the trace-norm closure of the convex hull of the states $\tau_v = \mathcal{E}_v(\rho_0)/\Pr_{\rho_0}(v)$, over all words $v$ with $\Pr_{\rho_0}(v) > 0$, the empty word included. If $\tau \in K$ and $\Pr_\tau(q) = 1$ for an outcome $q$, then $\Pr_\tau(q, u) = \Pr_\tau(u)$ for every word $u$; so the state $\sigma$ of the proposition is not in $K$. Indeed $\Pr_{\tau_v}(u) = \Pr_{e/3}(v, u)/\Pr_{e/3}(v) = \Pr_{\rho_v}(u)$ for the normalized state $\rho_v$ of $J$ after $v$, so on a finite convex combination $\tau$ of the $\tau_v$ the table is that of the same combination $\rho(\tau)$ of the $\rho_v$. The elements $\theta(w)$ of (3.2) span $J$ (proof of Proposition 8.4), so finitely many word probabilities determine $\rho(\tau)$ linearly, and each $\Pr_\tau(u) = \operatorname{Tr} \mathcal{E}_u(\tau)$ is 1-Lipschitz in trace norm (proof of Proposition 6.20). Hence, if $\tau_n \to \tau$ with every $\tau_n$ in the convex hull, the states $\rho(\tau_n)$ converge to a state $\rho$ of $J$ with $\Pr_\rho(u) = \Pr_\tau(u)$ for all $u$. If $\Pr_\tau(q) = 1$, then $U_q \rho = \rho$ by Lemma 3.1, so $\Pr_\rho(q, u) = \Pr_\rho(u)$.

# 7 Five tests, every four of them quantum

For a whole family of tests, four at a time is not enough. Theorem 7.2 gives five tests of atoms, no two orthogonal, every four of which have a sequential projective representation while the five together have none. Five always suffices (Corollary 7.5). Words of length at most eleven already exclude the five tests, by the length-seven argument applied to four atoms obtained from them by a reflection; this bound is proved in the companion paper (section "The reflected quadruple and its budget") and machine-checked (Appendix G).

## 7.1 Coordinates

Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O}\}$ and let $H_3(\mathbb{K})$ be the Hermitian $3 \times 3$ matrices over $\mathbb{K}$ with $x \circ y = \frac{1}{2}(xy + yx)$; for $\mathbb{K} = \mathbb{O}$ this is $J$. The Cayley–Dickson units are $i = e_1$, $j = e_2$, $\ell = e_4$ of Appendix A; the products of $i, j, \ell$ give the remaining units, so the only subalgebra of $\mathbb{O}$ containing $1, i, j, \ell$ is $\mathbb{O}$. For $x \in \mathbb{K}$ and matrix places $a \neq b$ let $F_{ab}(x)$ be the matrix with $x$ in place $(a, b)$, $\bar{x}$ in place $(b, a)$ and zeros elsewhere, so $F_{ba}(x) = F_{ab}(\bar{x})$, and let $E_{aa}$ be the diagonal matrix units. Matrix places are written $a, b, c$ here and in Appendix E, since $i$ and $j$ denote octonion units. (In the slot notation of Appendix A.3, $F_{23}(x) = F_1(x)$, $F_{31}(x) = F_2(x)$ and $F_{12}(x) = F_3(x)$.) Then $H_3(\mathbb{K})$ is the direct sum of $\mathbb{R}E_{11}$, $\mathbb{R}E_{22}$, $\mathbb{R}E_{33}$ and the three spaces $F_{ab}(\mathbb{K})$, and for distinct $a, b, c$

$$F_{ab}(x) \circ F_{bc}(y) = \tfrac{1}{2} F_{ac}(xy). \tag{7.1}$$

For a column $v \in \mathbb{K}^3$ whose entries lie in one associative subalgebra of $\mathbb{K}$ (automatic if they involve at most two imaginary units, by Artin's theorem) put $p(v) = vv^\dagger/\|v\|^2$; it is a primitive idempotent, and its $(a, b)$ entry is $v_a \bar{v}_b/\|v\|^2$. For two such columns whose entries lie jointly in one associative subalgebra, the computation takes place in a copy of $H_3(\mathbb{H})$, where $\langle x, y \rangle = \operatorname{Re} \operatorname{Tr}(xy)$, and gives the transition probability

$$\operatorname{tr}\big(p(u) \circ p(v)\big) = \frac{|u^\dagger v|^2}{\|u\|^2 \|v\|^2}. \tag{7.2}$$

**Lemma 7.1** (Matrices over a subalgebra)**.** *Let $B$ be a unital subalgebra of $H_3(\mathbb{K})$ that contains $E_{11}, E_{22}, E_{33}$, $F_{ac}(1)$ and $F_{bc}(1)$, where $\{a, b, c\} = \{1, 2, 3\}$. Let $\mathbb{F} = \{x \in \mathbb{K} : F_{12}(x) \in B\}$. Then $\mathbb{F}$ is a unital subalgebra of $\mathbb{K}$ closed under conjugation, and $B = H_3(\mathbb{F})$. If $B = \mathrm{Jord}\langle e, X\rangle$, then $\mathbb{F}$ is the smallest unital subalgebra of $\mathbb{K}$ closed under conjugation that contains the off-diagonal entries of all elements of $X$.*

*Formalized.* `lem_five_coord`, `lem_five_coord_gen`, with subalgebras of $H_3(\mathbb{K})$ taken inside $J$.

The proof is in Appendix E.

Two atoms $a \neq b$ of $H_3(\mathbb{K})$ that lie under the same rank-two idempotent $f$, that is, in the Peirce space $J_0(q)$ of the primitive idempotent $q = e - f$, determine $f$ (proof in Appendix E):

$$f = \frac{a + b - 2\,a \circ b}{1 - \mathrm{tr}(a \circ b)}. \tag{7.3}$$

The atoms $p \leq f$ are called the atoms *on the line* $f$ (the projective line of all atoms orthogonal to $q$, in the sense of[66]).

## 7.2 A five-atom configuration

**Theorem 7.2** (Five tests, every four quantum)**.** *Consider the five atoms of $J$*

$$\mathsf{F} = \big(p_0, p_1, p_2, p_3, p_4\big) = \big(E_{11},\ p(1,1,0),\ p(1,i,1),\ p(1,j,0),\ p(1,\ell,0)\big). \tag{7.4}$$

*Formalized.* (i)–(iii): `pentadF_isAtom`, `pentadF_generatesAlbert`, `pentadDrop_generated_finrank`; (a) `thm_pentad_a`; (b) `LengthSeven.Uniform.pentad_b_algebraic`, `FiveTest.thm_pentad_b`.

*(i) No two are orthogonal:* $\mathrm{tr}(p_2 \circ p_k) = \frac{1}{3}$ *for* $k \neq 2$*, and the other six transition probabilities equal* $\frac{1}{2}$*.*

*(ii)* $\mathrm{Jord}\langle e, p_0, \ldots, p_4\rangle = J$*.*

*(iii) Each four-element subfamily* $\mathsf{F} \setminus \{p_k\}$ *lies in a unital subalgebra of $J$ isomorphic to $H_3(\mathbb{H})$. The algebras generated by* $\mathsf{F} \setminus \{p_k\}$*,* $k = 0, \ldots, 4$*, have dimensions* $15, 15, 6, 15, 15$*.*

*For the tests of these atoms and model class (I) (Definition 3.4):*

*(a) for every state $\rho$ of $J$ and every $k$, the four tests $\mathsf{F} \setminus \{p_k\}$ have a sequential projective representation on $\mathbb{C}^6$ with a density-matrix state and one projection of rank two per test;*

*(b) for no state of $J$ do the five tests have a sequential projective representation, over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$, in any dimension, with projections of any rank and any state of the representation.*

*Proof. (i)* Each pair of columns involves at most two of $i, j, \ell$, so (7.2) applies. For instance $|(1,1,0)^\dagger(1,i,1)|^2 = |1+i|^2 = 2$ gives $\frac{2}{2\cdot 3} = \frac{1}{3}$, and $|(1,j,0)^\dagger(1,\ell,0)|^2 = |1 - j\ell|^2 = 2$ gives $\frac{2}{4} = \frac{1}{2}$; the other pairs are alike.

*(ii)* Explicitly,

$$p_1 = \tfrac{1}{2}(E_{11} + E_{22}) + \tfrac{1}{2}F_{12}(1), \qquad p_2 = \tfrac{1}{3}\big(e + F_{12}(-i) + F_{13}(1) + F_{23}(i)\big),$$
$$p_3 = \tfrac{1}{2}(E_{11} + E_{22}) + \tfrac{1}{2}F_{12}(-j), \quad p_4 = \tfrac{1}{2}(E_{11} + E_{22}) + \tfrac{1}{2}F_{12}(-\ell).$$

[66] Günaydin, Piron and Ruegg, *Moufang plane and octonionic quantum mechanics*, note 29.

Let $B$ be the generated algebra. It contains $E_{11} = p_0$, hence $U_{e-E_{11}} p_1 = \frac{1}{2} E_{22}$, hence $E_{22}$ and $E_{33} = e - E_{11} - E_{22}$. Peirce components with respect to this frame give $F_{12}(1)$ (from $p_1$) and $F_{13}(1)$ (from $p_2$). Lemma 7.1, with $c = 1$, gives $B = H_3(\mathbb{F})$, where $\mathbb{F}$ contains the entries $i, j, \ell$. So $\mathbb{F} = \mathbb{O}$ and $B = J$.

*(iii)* The proof is in Appendix E.

*(a)* Let $Q \cong H_3(\mathbb{H})$ contain the four atoms, $\psi : H_3(\mathbb{H}) \to Q$ an isomorphism, and $\iota : H_3(\mathbb{H}) \to H_6(\mathbb{C})$ the symplectic embedding, a unital Jordan monomorphism. An atom of $J$ in $Q$ is primitive in $Q$, so $\iota\psi^{-1}$ sends it to a rank-two projection. Restrict $\rho$ to $Q$, transport it along $\iota\psi^{-1}$, and extend it to a density matrix as in Section 4.4; then every word has its octonionic probability, since $\iota\psi^{-1}$ preserves quadratic maps.

*(b)* By (ii) and Theorem 4.1. □

By (ii) and (iii), no proper subfamily of $\mathsf{F}$ generates $J$.

### 7.3 Five tests always suffice

A family $S$ of primitive idempotents is *irredundant* if no proper subfamily generates the same unital subalgebra. Every subfamily of an irredundant family is irredundant (a redundancy in the subfamily is one in $S$). For a Euclidean Jordan algebra $A$ of rank three let $k_{\max}(A)$ be the largest size of an irredundant family of primitive idempotents that generates $A$.

**Lemma 7.3** (Reduction to irredundant families)**.** *Let $h$ be the least number such that a set of atoms of $J$ lies in a proper unital subalgebra whenever each of its subsets of at most $h$ elements does. Then $h = k_{\max}(J)$.*

*Formalized.* `lem_five_helly`.

*Proof.* If $S$ lies in no proper unital subalgebra, it generates $J$. The generated algebra is spanned by finitely many products of elements of $S$, so a finite subset generates $J$, and removing members while the remainder still generates gives an irredundant generating subset, of size at most $k_{\max}(J)$; so $h \le k_{\max}(J)$. An irredundant generating family of size $k_{\max}(J)$ has all its proper subsets in proper subalgebras; so $h \ge k_{\max}(J)$. □

**Proposition 7.4** (Largest irredundant generating families)**.**

*Formalized.* `prop_arity`.

$$\big(k_{\max}(H_3(\mathbb{R})),\ k_{\max}(H_3(\mathbb{C})),\ k_{\max}(H_3(\mathbb{H})),\ k_{\max}(J)\big) = (3, 4, 4, 5).$$

The lower bounds are explicit families. The proof of the upper bounds, in Appendix E, has two ingredients. First, every proper unital subalgebra of $H_3(\mathbb{K})$ lies in a *line subalgebra* $\mathbb{R}q \oplus J_0(q)$, for a primitive idempotent $q$, or in a copy of $H_3(\mathbb{K}')$, where $\mathbb{K}'$ is the next smaller of $\mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O}$; this is Lemma E.1, proved from the classification. Second, one atom off a line together with several atoms on it generates $H_3(\mathbb{F})$, where $\mathbb{F}$ is generated by one off-diagonal entry for each atom on the line. An induction along $\mathbb{R}, \mathbb{C}, \mathbb{H}, \mathbb{O}$ combines the two.

In the language of combinatorial convexity, five is the *Helly number* of the family of proper subalgebras of $J$, the number $h$ of Lemma 7.3. The same argument with Proposition 7.4 gives the Helly numbers 3, 4, 4 for the proper subalgebras of $H_3(\mathbb{R})$, $H_3(\mathbb{C})$, $H_3(\mathbb{H})$.

**Corollary 7.5** (Helly numbers)**.** *A set $S$ of atoms of $J$, finite or infinite, lies in a proper (equivalently, special) unital subalgebra of $J$ if and only if each subset of at most five members does, and five cannot be lowered. For model class (I): a finite family of tests of atoms of $J$ has a sequential projective representation (from some, equivalently every, state of $J$) if and only if each subfamily of at most five of its tests does.*

*Narrower in Lean.* First sentence: `cor_helly_geometric`, `cor_helly_special`, and `cor_helly_sharp` for five. Second sentence: `AlbertChallenge.Bridge.cor_helly_classI`, for spaces in Lean's lowest universe, algebraic models admitted, subfamilies indexed by injections; in algebraic-model form, `cor_helly_operational_iff_unconditional`.

*Proof.* Lemma 7.3 and Proposition 7.4. A proper unital subalgebra has dimension below 27, so it is special by the classification. The statement about representations is Theorem 4.1, which shows that existence depends only on whether the generated algebra is $J$. □

**Local and global.** All five four-element subfamilies of $\mathsf{F}$ have $\mathcal{A} = 0$, as for five atoms in one copy of $H_3(\mathbb{H})$, so whether five atoms generate $J$ is not a function of the four-point invariants of their subfamilies. In the marginal problem, locally consistent data on a cover extend for all data exactly when the cover is acyclic[67], a property of the cover; here the number five is a property of the algebra $J$. For joint measurability of sharp observables in one Hilbert space the analogous number is two (pairwise commutation), while unsharp observables can be pairwise but not triplewise jointly measurable[68].

# 8 Where the projection postulate enters

The main theorem compares the octonionic table with model class (I) (Definition 3.4). Letting the hidden quantum system answer each test with more general instruments shows which part of the projection postulate the theorem uses.

For one test of an atom $p$, the state after the first outcome is a multiple of $p$ or lies in $J_0(p)$, later tests of $p$ change nothing, and a two-state hidden Markov model suffices.

## 8.1 A finite realization with non-Lüders instruments

**Theorem 8.1** (A finite completely positive realization of ideal-test statistics with non-Lüders instruments)**.** *Let $p_1, \dots, p_n$ be atoms of $J$ in any configuration and $\rho$ any state of $J$. The model class is (II) (Definition 3.6). There are a density matrix $\sigma$ on $\mathbb{C}^{17n+1}$ and, for each test $j$, completely positive maps $\mathcal{E}_{\mathrm{yes}_j}$, $\mathcal{E}_{\mathrm{no}_j}$ on $M_{17n+1}(\mathbb{C})$ with trace-preserving sum, all given by real matrices, such that*

$$\operatorname{Tr}\bigl(\mathcal{E}_{q_m} \cdots \mathcal{E}_{q_1} \sigma\bigr) = \Pr_\rho(q_1, \dots, q_m)$$

*for every word. For four tests the dimension is* 69.

*Formalized.* `NativeAnalysis.exists_nativeGeneral_real_simulation`; the dimension $17n + 1$: `NativeAnalysis.nativeTaggedBlock_card`.

[67] N. N. Vorob'ev, Consistent families of measures and their extensions, *Theory Probab. Appl.* **7** (1962) 147–163. doi:10.1137/1107014; S. Abramsky and A. Brandenburger, The sheaf-theoretic structure of non-locality and contextuality, *New J. Phys.* **13** (2011) 113036. doi:10.1088/1367-2630/13/11/113036. arXiv:1102.0264.

[68] T. Heinosaari, D. Reitzner and P. Stano, Notes on joint measurability of quantum observables, *Found. Phys.* **38** (2008) 1133–1147. doi:10.1007/s10701-008-9256-7. arXiv:0811.0783.

In the language of Budroni, Fagundes and Kleinmann[69], this is a quantum finite-state machine of dimension $17n + 1$ with general instruments. Because it uses one fixed instrument per test, it reproduces the statistics under every rule for choosing the tests, random or adaptive. Tests of idempotents of rank two are tests of atoms with relabelled outcomes, and ranks 0 and 3 are trivial, so every finite family of tests on $J$ has a finite completely positive realization.

**The idea.**

(i) *A "yes" resets.* $U_{p_j} x = \langle p_j, x \rangle\, p_j$, so after $\text{yes}_j$ the Albert state is a known multiple of $p_j$; the memory stores one number.

(ii) *A "no" lands in a spin factor.* $U_{e-p_j}$ is the orthogonal projection onto the Peirce space $J_0(p_j) \cong H_2(\mathbb{O})$, a ten-dimensional spin factor whose cone is the forward light cone of Minkowski space $\mathbb{R}^{1,9}$. Spin factors are special: they embed in real $16 \times 16$ matrices through a Clifford representation, a standard fact[70].

(iii) *Passing between Peirce spaces is a Lorentz similitude unless the atoms are orthogonal.* The restriction of $U_{e-p_j}$ to $J_0(p_i)$ multiplies the Minkowski quadratic form by the transition probability $o_{ij} = \operatorname{tr}(p_i \circ p_j)$. If $o_{ij} > 0$ it is carried by a single $16 \times 16$ spinor matrix; if $o_{ij} = 0$ it has rank one and is realized by measuring and preparing.

The memory has a basis vector $|\text{init}\rangle$ for the initial state, one vector $|P_k\rangle$ per atom $p_k$ (a weight $t$ on it encodes the state $t\, p_k$ after a "yes"), and one 16-dimensional block per Peirce space $J_0(p_k)$, holding a spinor-level copy of the state there: $1 + n + 16n$ dimensions. The proof is in Appendix D.

The realization does not embed the full state-and-effect model of $J$ in quantum theory, preserving all outcome probabilities; such an embedding into finite-dimensional quantum theory is impossible[71]. This obstruction concerns both states and effects, not a map of states alone. It encodes only the states that ideal tests reach after at least one test, and these lie in the special pieces $\mathbb{R}p_j$ and $J_0(p_j)$.

**Corollary 8.2** (Omitting the initial memory state)**.** *In Theorem 8.1, suppose that for one of the tests $i$ the initial state has the form $\rho = t p_i + y$, with $t \geq 0$ and $y \in J_0(p_i)_+$; equivalently, its component in $J_{1/2}(p_i)$ is zero. Then the realization can be made on $\mathbb{C}^{17n}$, with real data. In particular this applies to $\rho = e/3$, giving dimension 68 for four tests.*

*Formalized.* `NativeAnalysis.exists_nativeNoHalf_real_simulation`; 68 for four tests: `NativeAnalysis.exists_nativeFourIdempotent_mixed_instrument_68`.

*Proof.* Delete only the initial ray $\mathbb{C}|\text{init}\rangle$ from the construction of Appendix D. The remaining reset rays $\mathbb{C}|P_k\rangle$ and spinor blocks are invariant under every branch, and

[69] C. Budroni, G. Fagundes and M. Kleinmann, Memory cost of temporal correlations, *New J. Phys.* **21** (2019) 093018. doi:10.1088/1367-2630/ab3cb4. arXiv:1902.06517.

[70] Jordan, von Neumann and Wigner, *On an algebraic generalization*, note 4; Faraut and Korányi, *Analysis on Symmetric Cones*, note 16.

[71] A. J. P. Garner and M. P. Müller, Characterization of the probabilistic models that can be embedded in quantum theory, arXiv:2004.06136 (2020); M. P. Müller and A. J. P. Garner, Testing quantum theory by generalizing noncontextuality, *Phys. Rev. X* **13** (2023) 041001. doi:10.1103/PhysRevX.13.041001. arXiv:2112.09719.

the restrictions remain completely positive with trace-preserving sum. Initialize them with $\sigma = t|P_i\rangle\langle P_i| \oplus \mathrm{enc}_i(y)$, with $\mathrm{enc}_i$ the encoding of that appendix and zero on the other blocks. The encoding preserves trace, so $\operatorname{Tr}\sigma = t + \operatorname{tr} y = 1$. Each branch advances both summands to the encoding of the corresponding Albert update, by the same block formulas as in that appendix; linearity handles the first test and its induction handles every later word. For $\rho = e/3$, explicitly, $t = 1/3$ and $\mathrm{enc}_i((e - p_i)/3) = I_{16}/24$. □

*Remark* 8.3 (The realization is not projective). For atoms that generate $J$, the effect $\mathcal{E}^*_{\mathrm{yes}_j}(1)$ is not a projection for some $j$: on the block of a Peirce space $J_0(p_i)$, $i \neq j$, with $0 < o_{ij} < 1$ its spectrum is $\{0, 1 - o_{ij}\}$, each eigenvalue eight times (Appendix D). Such a pair exists whenever the atoms generate $J$, since atoms with all $o_{ij} \in \{0, 1\}$ lie in the algebra of one Jordan frame. The output maps are not Lüders maps of their effects, so the realization is not in class (I).

**Proposition 8.4** (Lower bound). *Let $p_1, \ldots, p_n$ be atoms that generate $J$, and $\rho = e/3$. Every completely positive realization of their tests (class (II)) has dimension $d \geq 6$. For four such atoms the least dimension $d_{\min}$ satisfies $6 \leq d_{\min} \leq 68$.*

*Formalized.* `native_CP_dimension_at_least_six`; the upper bound 68: `NativeAnalysis.exists_nativeFourIdempotent_mixed_instrument_68`.

*Proof.* In a realization of dimension $d$, the Hankel matrix $H_{u,v} = \Pr(uv) = \operatorname{Tr}\big(\mathcal{E}_v\mathcal{E}_u\sigma\big) = \langle\mathcal{E}^*_v(1), \mathcal{E}_u\sigma\rangle_{\mathrm{HS}}$ factors through the real space of Hermitian $d \times d$ matrices, so its rank is at most $d^2$. Bounding the dimension of a realization by the rank of its Hankel matrix is the basic step of the realization theory of recognizable series (Schützenberger[72], Fliess[73]).

On the Albert side, $U_q$ is self-adjoint, so with $\theta(w)$ as in (3.2)

$$H_{u,v} = \operatorname{tr}\big(U_{(v)}U_{(u)}\tfrac{e}{3}\big) = \tfrac{1}{3}\langle\theta(v), \theta(u^{\mathrm{rev}})\rangle,$$

where $U_{(u)} = U_{u_k} \cdots U_{u_1}$ is the composite update along $u = (u_1, \ldots, u_k)$ and $u^{\mathrm{rev}}$ is $u$ reversed. So $H$ is the Gram matrix of the $\theta(w)$ with its rows indexed by reversed words, and its rank is $\dim \operatorname{span}\{\theta(w)\}$. This span contains $e$ and each $p_i$, and it is invariant under $U_{p_i}$ and $U_{e-p_i}$, hence under the reflections $U_{2p_i-e} = 2U_{p_i} + 2U_{e-p_i} - \mathrm{id}$. It therefore contains every orbit point $q_\alpha$, and by Lemma 4.7 it is $J$; that lemma's proof uses only (4.2), which holds because atoms of $J$ are primitive in $B = J$. So $d^2 \geq 27$, that is $d \geq 6$. The upper bound is Corollary 8.2. □

## 8.2 Sharp effects with arbitrary post-measurement channels

A realization can also have *sharp* effects, one fixed projection per test, if the update after each outcome is an arbitrary channel.

The factorisation of an instrument into a Lüders instrument followed by outcome-dependent channels, and the Naimark dilation of an effect to a projection,

[72] M.-P. Schützenberger, On the definition of a family of automata, *Inf. Control* **4** (1961) 245–270. doi:10.1016/S0019-9958(61)80020-X.

[73] M. Fliess, Matrices de Hankel, *J. Math. Pures Appl.* (9) **53** (1974) 197–222.

go back to Davies and Lewis[74] and Ozawa[75]. What is needed here is that the dilation can be made for all tests at once, on one space, with one fixed projection and two fixed channels per test, so that it reproduces *sequences* of tests without a growing record of the past; the short proofs follow. In fixed dimension, instruments of the form "projective measurement, then a channel" are a strict subclass of all instruments (Khandelwal and Tavakoli[76]); in doubled dimension they reproduce the sequential statistics of any two-outcome realization.

For a two-outcome instrument $(\mathcal{E}_+, \mathcal{E}_-)$ on $M_d(\mathbb{C})$ write $A = \mathcal{E}_+^*(1)$, so $0 \le A \le 1$ and $\mathcal{E}_-^*(1) = 1 - A$.

**Lemma 8.5** (Factoring through the square root of the effect)**.** *Let $\mathcal{E}$ be a completely positive map on $M_d(\mathbb{C})$ with $\mathcal{E}^*(1) = A$, let $\Pi$ be the support projection of $A$, $A^{-1/2}$ the inverse of $\sqrt{A}$ on* $\operatorname{ran}\Pi$ *(extended by 0), and $\tau_0$ a density matrix. Then*

$$\mathcal{F}(Y) = \mathcal{E}\big(A^{-1/2} Y A^{-1/2}\big) + \operatorname{Tr}\big((1-\Pi)Y\big)\,\tau_0$$

*is a channel and $\mathcal{E}(X) = \mathcal{F}(\sqrt{A}X\sqrt{A})$ for all $X$.*

*Narrower in Lean.* `exists_kraus_channel_factorization`. Existence of the channel, in Kraus form; the formula for $\mathcal{F}$ is not stated.

*Proof.* Write $\mathcal{E}(X) = \sum_k K_k X K_k^*$ with $\sum_k K_k^* K_k = A$. Each $K_k$ vanishes on $\ker A$, since $\|K_k x\|^2 \le \langle x, Ax\rangle$. So $L_k = K_k A^{-1/2}$ satisfies $L_k\sqrt{A} = K_k\Pi = K_k$ and $\sum_k L_k^* L_k = A^{-1/2} A A^{-1/2} = \Pi$. Hence $\mathcal{F}$ is completely positive with $\mathcal{F}^*(1) = \Pi + (1-\Pi) = 1$, and $\mathcal{F}(\sqrt{A}X\sqrt{A}) = \sum_k K_k X K_k^* = \mathcal{E}(X)$ because $(1-\Pi)\sqrt{A} = 0$. □

**Proposition 8.6** (Sharp effects with arbitrary channels)**.** *Let $\sigma$ and two-outcome instruments $(\mathcal{E}_{i,+}, \mathcal{E}_{i,-})$, $i = 1, \dots, n$, on $M_d(\mathbb{C})$ form a completely positive realization (Definition 3.6) of some process, and let $A_i = \mathcal{E}_{i,+}^*(1)$. Put $\mathcal{K}_2 = \mathbb{C}^d \oplus \mathbb{C}^d$ and $\iota x = (x, 0)$. For each $i$ there are an orthogonal projection $P_i$ on $\mathcal{K}_2$ and channels $\Lambda_{i,\pm}$ on $B(\mathcal{K}_2)$ such that the instruments*

$$\widetilde{\mathcal{E}}_{i,+}(Z) = \Lambda_{i,+}(P_i Z P_i), \qquad \widetilde{\mathcal{E}}_{i,-}(Z) = \Lambda_{i,-}\big((1-P_i)Z(1-P_i)\big)$$

*have the following properties.*

*Narrower in Lean.* `exists_supported_sharpDilation`. The last sentence, on real Kraus operators, is not part of this declaration; for Corollary 8.7 it is proved.

*(a)* $\iota^* P_i \iota = A_i$*, the effect of $\widetilde{\mathcal{E}}_{i,+}$ is $P_i$, and $\widetilde{\mathcal{E}}_{i,\pm}(\iota X \iota^*) = \iota \mathcal{E}_{i,\pm}(X)\iota^*$ for all $X \in M_d(\mathbb{C})$.*

*(b) Every output of every $\widetilde{\mathcal{E}}_{i,\pm}$ lies in $\iota M_d(\mathbb{C})\iota^*$.*

*(c) $\iota\sigma\iota^*$ and the instruments $(\widetilde{\mathcal{E}}_{i,+}, \widetilde{\mathcal{E}}_{i,-})$ form a completely positive realization of the same process on $\mathbb{C}^{2d}$.*

*(d) If each $\mathcal{E}_{i,+}$ maps positive matrices to matrices supported in the eigenvalue-1 space of $A_i$, and each $\mathcal{E}_{i,-}$ into* $\ker A_i$*, the channels can be chosen so that the new instruments are repeatable on all states: $\widetilde{\mathcal{E}}_{i,+}(Z) = P_i \widetilde{\mathcal{E}}_{i,+}(Z) P_i$ and $\widetilde{\mathcal{E}}_{i,-}(Z) = (1-P_i)\widetilde{\mathcal{E}}_{i,-}(Z)(1-P_i)$ for $Z \ge 0$.*

[74] E. B. Davies and J. T. Lewis, An operational approach to quantum probability, *Commun. Math. Phys.* **17** (1970) 239–260.

[75] M. Ozawa, Quantum measuring processes of continuous observables, *J. Math. Phys.* **25** (1984) 79–87.

[76] S. Khandelwal and A. Tavakoli, Simulating quantum instruments with projective measurements and quantum post-processing, *Phys. Rev. Lett.* **135** (2025) 040202. arXiv:2503.00956.

*If $\sigma$ and all Kraus operators are real matrices, the projections $P_i$ and the channels $\Lambda_{i,\pm}$ can be chosen with real Kraus operators.*

*Proof.* Fix $i$ and drop it from the notation. If $A = 0$, put $P = 0$, let $\Lambda_+$ be any channel and $\Lambda_-(Z) = \iota\mathcal{E}_-(\iota^* Z \iota)\iota^* + \operatorname{Tr}((1 - \iota\iota^*)Z)\,\iota\tau_0\iota^*$; if $A = 1$, exchange the roles of $+$ and $-$. All claims are then immediate, since $\mathcal{E}_+ = 0$ when $A = 0$. From now on $A \neq 0, 1$.

*The projection.* Since $\sqrt{A}$ and $\sqrt{1-A}$ commute,

$$U = \begin{pmatrix} \sqrt{A} & \sqrt{1-A} \\ \sqrt{1-A} & -\sqrt{A} \end{pmatrix}$$

is a self-adjoint unitary on $\mathcal{K}_2$. Let $\iota' x = (0, x)$, $\iota_+ = U\iota$ and $\iota_- = U\iota'$. These are isometries with orthogonal ranges spanning $\mathcal{K}_2$, so $P = \iota_+\iota_+^*$ is a projection, $1 - P = \iota_-\iota_-^*$, and $\iota^* P \iota = A$. Moreover

$$P\iota = U\,\iota\iota^*\,U\iota = \iota_+\sqrt{A}, \qquad (1-P)\iota = U\,\iota'\iota'^*\,U\iota = \iota_-\sqrt{1-A}. \tag{8.1}$$

*The channels.* By Lemma 8.5 choose channels $\mathcal{F}_\pm$ with $\mathcal{E}_+(X) = \mathcal{F}_+(\sqrt{A}X\sqrt{A})$ and $\mathcal{E}_-(X) = \mathcal{F}_-(\sqrt{1-A}X\sqrt{1-A})$, fix density matrices $\tau_\pm$ on $\mathcal{K}_2$, and put

$$\Lambda_+(Z) = \iota\mathcal{F}_+(\iota_+^* Z \iota_+)\iota^* + \operatorname{Tr}\big((1-P)Z\big)\,\tau_+,$$
$$\Lambda_-(Z) = \iota\mathcal{F}_-(\iota_-^* Z \iota_-)\iota^* + \operatorname{Tr}(PZ)\,\tau_-.$$

Both are completely positive, and $\operatorname{Tr}\Lambda_+(Z) = \operatorname{Tr}(\iota_+\iota_+^* Z) + \operatorname{Tr}((1-P)Z) = \operatorname{Tr} Z$; likewise for $\Lambda_-$.

*(a)–(c).* Since $\iota_+^* P = \iota_+^*$ and $(1-P)P = 0$, it follows that $\widetilde{\mathcal{E}}_+(Z) = \iota\mathcal{F}_+(\iota_+^* Z \iota_+)\iota^*$, and likewise $\widetilde{\mathcal{E}}_-(Z) = \iota\mathcal{F}_-(\iota_-^* Z \iota_-)\iota^*$; this is (b). The effects are $P$ and $1 - P$. By (8.1), $\iota_+^*\iota X \iota^*\iota_+ = \iota_+^* P \iota X \iota^* P \iota_+ = \sqrt{A}X\sqrt{A}$, so $\widetilde{\mathcal{E}}_+(\iota X \iota^*) = \iota\mathcal{E}_+(X)\iota^*$, and the same with $\iota_-$ for the outcome $-$. This is (a). Induction on the word gives $\widetilde{\mathcal{E}}_w\,\iota\sigma\iota^* = \iota(\mathcal{E}_w\sigma)\iota^*$; taking traces gives (c).

*(d).* Let $E_1$ and $E_0$ be the eigenvalue-1 space and the kernel of $A$; both are nonzero, since $\mathcal{E}_+ \neq 0 \neq \mathcal{E}_-$. Take the state $\tau_0$ of $\mathcal{F}_+$ in $E_1$ and that of $\mathcal{F}_-$ in $E_0$. Then $\mathcal{F}_+$ maps positive matrices into matrices supported in $E_1$, and $\mathcal{F}_-$ into $E_0$. By (8.1), $(1-P)\iota x = \iota_-\sqrt{1-A}\,x = 0$ for $x \in E_1$ and $P\iota x = 0$ for $x \in E_0$, so $\iota E_1 \subseteq \operatorname{ran} P$ and $\iota E_0 \subseteq \operatorname{ran}(1-P)$. With (b) this gives (d).

The reality statement holds, choosing $\tau_0, \tau_\pm$ real, because $U$, $\iota_+$, $\iota_-$, $P$, $A^{-1/2}$ and the $L_k$ are obtained from the given matrices by real operations and the real spectral theorem. □

Property (b) is what makes the construction compose: whichever test comes next finds the state in $\iota M_d(\mathbb{C})\iota^*$, where its own dilation expects it.

**Corollary 8.7** (Ideal-test statistics with sharp effects)**.** *Let $p_1, \dots, p_n$ be atoms of $J$ in any configuration and $\rho$ any state. There are a density matrix on $\mathbb{C}^{2(17n+1)}$ and, for each test $j$, one orthogonal projection $P_j$ and two channels $\Lambda_{j,\pm}$, all real matrices, such that the instruments $Z \mapsto \Lambda_{j,+}(P_j Z P_j)$ and $Z \mapsto \Lambda_{j,-}((1-P_j)Z(1-P_j))$ reproduce $\Pr_\rho(w)$ for every word and are repeatable on all states. For four tests the dimension is* 138*. Under the initial-state hypothesis of Corollary 8.2, the dimension can instead be* $34n$*, hence* 136 *for four tests, in particular from $e/3$.*

*Formalized.* `NativeAnalysis.exists_nativeSharp_instrument_2_17n_add_one`, `NativeAnalysis.exists_nativeNoHalf_sharp_instrument_34n`.

*Proof.* Apply Proposition 8.6 to the real realization of Theorem 8.1; the hypothesis of (d) follows from the block description of Appendix D. Every instrument there first discards coherences between blocks, so $A_j$ is block diagonal. Every output of $\mathcal{E}_{\mathrm{yes}_j}$ is a multiple of $|P_j\rangle\langle P_j|$, and $A_j|P_j\rangle = |P_j\rangle$ because the test of $p_j$ is certain to answer "yes" on the block of $p_j$. Every output of $\mathcal{E}_{\mathrm{no}_j}$ lies in the block of $J_0(p_j)$, where $A_j = 0$. The same properties hold after deleting the initial ray as in Corollary 8.2, so doubling that realization gives $34n$. □

In the companion paper (section "Repeatability and a smaller sharp dilation") the added space has dimension $\max_i \operatorname{rank}(A_i - A_i^2)$ instead of $d$, which lowers these dimensions to $26n - 7$ for $n \geq 1$, and to $26n - 9$ under the initial-state hypothesis (97 and 95 for four tests).

*Remark* 8.8 (What fails is ideality)*.* Measure the instruments of Corollary 8.7 against Definition 3.2.

*Sharpness (Sh) holds.* The faces are the states supported in $\operatorname{ran} P_j$ and in $\operatorname{ran}(1-P_j)$; an effect vanishing on both has $P_j f P_j = (1-P_j) f (1-P_j) = 0$, hence $f = 0$.

*Ideality (Id) fails.* Let $a = \operatorname{tr}(p_j \circ \rho)$ with $0 < a < 1$ (for $\rho = e/3$, $a = \frac{1}{3}$), and $\xi = \iota_{j,+}|\mathrm{init}\rangle$, the image of the initial vector under the isometry $\iota_+$ of the proof of Proposition 8.6 for test $j$. Then $\xi \in \operatorname{ran} P_j$, so "yes" is certain on $|\xi\rangle\langle\xi|$. But $\iota_{j,+}^* \iota_{j,+} = 1$, Lemma 8.5 and $\mathcal{E}_{\mathrm{yes}_j}(|\mathrm{init}\rangle\langle\mathrm{init}|) = a|P_j\rangle\langle P_j|$ give

$$\widetilde{\mathcal{E}}_{j,+}\big(|\xi\rangle\langle\xi|\big) = \iota|P_j\rangle\langle P_j|\iota^* \neq |\xi\rangle\langle\xi|.$$

No realization with sharp effects can satisfy (Id) for every test: if one did, Theorem 3.3(b) would make all its instruments Lüders instruments, that is, a sequential projective representation, which Theorem 4.1 excludes when the atoms generate $J$.

**Example 8.9** (Two presses)**.** Take the instruments of Corollary 8.7 for $\mathsf{M}$ from $e/3$, on $\mathbb{C}^{68} \oplus \mathbb{C}^{68}$, with $\iota x = (x, 0)$ and $\iota' x = (0, x)$. The effects $A_A$ and $A_B$ are block diagonal, so on the reset ray $|P_C\rangle$ of Appendix D both act as $o_{AC} = o_{BC} = \frac{1}{3}$. Hence on the plane spanned by $\iota|P_C\rangle$ and $\iota'|P_C\rangle$ the projections $P_A$ and $P_B$ of Proposition 8.6 coincide, with range spanned by $\xi = \frac{1}{\sqrt{3}}\,\iota|P_C\rangle + \sqrt{\frac{2}{3}}\,\iota'|P_C\rangle$, and on $|\xi\rangle\langle\xi|$ both "yes" on $A$ and "yes" on $B$ are certain. Since $\xi = \iota_+|P_C\rangle$ for test $A$, the outcome "yes" on $A$ leaves $\iota|P_A\rangle\langle P_A|\iota^*$ (Lemma 8.5 and $\mathcal{E}_{\mathrm{yes}_A}(|P_C\rangle\langle P_C|) = \frac{1}{3}|P_A\rangle\langle P_A|$), on which "yes" on $B$ has probability $o_{AB} = \frac{1}{3}$. So a press of $A$ whose answer is certain lowers the probability of "yes" on $B$ from 1 to $\frac{1}{3}$. Every reachable state lies in $\iota M_{68}(\mathbb{C})\iota^*$, which does not contain $|\xi\rangle\langle\xi|$.

Sharp effects therefore do not exclude a finite-dimensional model, even with one fixed projection per test, time-independent channels and repeatability on all states. Given sharpness, which cannot be dropped (Example B.1), the hypothesis of the main theorem that does the work is ideality (Id), the Lüders update on both outcomes.

## 8.3 Complete measurements see only transition probabilities

If the tests are replaced by complete measurements of Jordan frames (Section 3.3), the exception disappears from the statistics.

**Proposition 8.10** (Complete measurements give a Markov chain)**.** *Let $F^1, \dots, F^m$ be Jordan frames of $J$ and $\rho$ a state, and perform complete measurements of these frames in any order, with repetitions. The probability of the outcomes $f_1, \dots, f_n$ (each $f_t$ an atom of the frame measured at step $t$) is*

$$\Pr(f_1, \dots, f_n) = \langle \rho, f_1 \rangle \prod_{t=1}^{n-1} \operatorname{tr}(f_t \circ f_{t+1}). \tag{8.2}$$

*Formalized.* `frameClassicalWordProbability_eq`, `frameMemory_card`; the comparison with $\mathbb{C}^3$: `mutuallyUnbiasedFrames_qutrit_word_statistics`.

*These statistics are those of a classical Markov chain on the atoms of the frames, and they have a hidden Markov model with $3m + 1$ states. For any $m \le 4$ pairwise mutually unbiased frames ($\operatorname{tr}(f \circ g) = \frac{1}{3}$ across frames) and $\rho = e/3$ they coincide with those of complete projective measurements of $m$ mutually unbiased bases of $\mathbb{C}^3$ from $I/3$.*

*Proof.* For an atom $f$, $U_f x = \langle f, x \rangle f$ by (3.1). Iterating,

$$U_{f_n} \cdots U_{f_1} \rho = \langle \rho, f_1 \rangle \prod_{t} \operatorname{tr}(f_t \circ f_{t+1})\, f_n,$$

whose trace is (8.2). The hidden state is the last outcome (or "initial"); from an atom $g$ a measurement of a frame $F$ moves to $f \in F$ with probability $\operatorname{tr}(g \circ f)$, and these sum to $\operatorname{tr}(g \circ e) = 1$. For mutually unbiased frames from $e/3$ the factors are $\frac{1}{3}$ at the first step, then 1 or 0 for a repeated frame (same or different outcome) and $\frac{1}{3}$ across frames. Mutually unbiased bases of $\mathbb{C}^3$ give the same factors, since $\operatorname{Tr}(P_\varphi P_\psi) = |\langle \varphi, \psi \rangle|^2$, and $\mathbb{C}^3$ has four of them. □

Frames can generate $J$, since any atom lies in a Jordan frame and four atoms can generate $J$. Their complete measurements still produce (8.2), which depends only on transition probabilities. In Hilbert space the same formula for sequences of rank-one complete projective measurements is the generalised Wigner formula; see Słomczyński and Szczepanek[77] and, for its origin, Wigner[78].

Two-outcome tests behave differently because their "no" update $U_{e-p}$ projects onto the rank-two Peirce space $J_0(p)$ and keeps the coherence inside it, whereas a complete measurement replaces that piece by a mixture of two atoms (compare Section 3.3).

[77] W. Słomczyński and A. Szczepanek, Quantum dynamical entropy, chaotic unitaries and complex Hadamard matrices, *IEEE Trans. Inf. Theory* **63** (2017) 7821–7831. doi:10.1109/TIT.2017.2751507. arXiv:1612.03363.

[78] E. P. Wigner, The problem of measurement, *Am. J. Phys.* **31** (1963) 6–15.

# 9 Related work

## 9.1 Reconstructions, composites and embeddings

Barnum, Müller and Ududec[79] list the three-level octonionic theory among the survivors of their single-system postulates and remove it by observability of energy, through the count $\dim F_4 = 52 > 27$. Barnum and Hilgert[80], Wilce[81] and Araki[82] reach Euclidean Jordan algebras by other routes and also keep $J$; Wilce notes that it is excluded only by a composite argument. Niestegge's conditional-probability postulates[83] allow the exceptional algebra, whose projection lattice has unique conditional probabilities and no Hilbert-space representation[84]. His transition-probability spaces with continuous reversible dynamics[85] again include the irreducible atomic JBW algebras, $J$ among them, together with an $E_6$-symmetric model that does not always provide post-measurement states; in every simple Euclidean Jordan algebra the transition probability is maximally homogeneous[86]. Van de Wetering[87] shows that spaces with a sequential product in the sense of Gudder and Greechie[88] are Euclidean Jordan algebras, $J$ among them. Westerbaan, Westerbaan and van de Wetering[89] split every normal sequential effect algebra into a complete Boolean algebra, a convex part and a purely almost-convex part.

Composites remove $J$. Barnum, Graydon and Wilce[90] prove that it has no composite with a nonclassical system when the composite is itself a Euclidean Jordan algebra carrying both factors' dynamics; locally tomographic composites

[79] Barnum, Müller and Ududec, *Higher-order interference and single-system postulates*, note 15.

[80] H. Barnum and J. Hilgert, Spectral properties of convex bodies, *J. Lie Theory* **30** (2020) 315–344. doi:10.5802/jolt.1118. arXiv:1904.03753 (preprint title: Strongly symmetric spectral convex bodies are Jordan algebra state spaces).

[81] Wilce, "Conjugates, filters and quantum mechanics", note 23.

[82] H. Araki, On a characterization of the state space of quantum mechanics, *Commun. Math. Phys.* **75** (1980) 1–24. doi:10.1007/BF01962588.

[83] G. Niestegge, Conditional probability, three-slit experiments, and the Jordan algebra structure of quantum mechanics, *Adv. Math. Phys.* **2012** (2012) 156573. doi:10.1155/2012/156573. arXiv:0912.0203; G. Niestegge, A simple and quantum-mechanically motivated characterization of the formally real Jordan algebras, *Proc. R. Soc. A* **476** (2020) 20190604. doi:10.1098/rspa.2019.0604. arXiv:1905.04189.

[84] G. Niestegge, A representation of quantum measurement in order-unit spaces, *Found. Phys.* **38** (2008) 783–795. doi:10.1007/s10701-008-9236-y. arXiv:1001.3633.

[85] G. Niestegge, The bare necessities of a physically reasonable mathematical model for quantum theory, arXiv:2607.24474 (2026).

[86] G. Niestegge, On the homogeneity of the quantum transition probability, arXiv:2601.01936 (2026).

[87] J. van de Wetering, Sequential product spaces are Jordan algebras, *J. Math. Phys.* **60** (2019) 062201. doi:10.1063/1.5093504. arXiv:1803.11139.

[88] Gudder and Greechie, "Sequential products on effect algebras", note 35.

[89] A. Westerbaan, B. Westerbaan and J. van de Wetering, The three types of normal sequential effect algebras, *Quantum* **4** (2020) 378. doi:10.22331/q-2020-12-24-378. arXiv:2004.12749.

[90] H. Barnum, M. A. Graydon and A. Wilce, Some nearly quantum theories, in *Proc. 12th Int. Workshop on Quantum Physics and Logic (QPL 2015)*, Electron. Proc. Theor. Comput. Sci. **195** (2015) 59–70. doi:10.4204/EPTCS.195.5. arXiv:1507.06278; H. Barnum, M. A. Graydon and A. Wilce, Composites and categories of Euclidean Jordan algebras, *Quantum* **4** (2020) 359. doi:10.22331/q-2020-11-08-359. arXiv:1606.09331.

force complex quantum theory (Barnum and Wilce[91]) or C*-algebras[92]. Local tomography of composites is also among the postulates of the reconstructions of Hardy[93], Masanes and Müller[94] and Chiribella, D'Ariano and Perinotti[95], which arrive at complex quantum theory.

Sorkin's third-order interference[96] vanishes in every Jordan-algebraic system, $J$ included (Ududec, Barnum and Emerson[97]; Niestegge[98]). Quaternionic quantum mechanics goes back to Finkelstein, Jauch, Schiminovich and Speiser[99], with a book-length treatment by Adler[100]; Peres's proposed test of complex against quaternionic quantum theory[101] separates two associative number systems. Real quantum theory reproduces every multipartite experiment of complex quantum theory when the parts share arbitrary real states (McKague, Mosca and Gisin[102]), but not in networks with independent sources (Renou et al.[103]).

The theorem nearest to the present theorem is that of Müller and Garner[104], announced by Garner and Müller[105]: an unrestricted finite-dimensional generalized probabilistic theory has an exact linear embedding into finite-dimensional quantum theory if and only if it is a special Euclidean Jordan algebra. They embed the whole state and effect spaces and preserve single-shot probabilities, with no update rule. Their criterion generalizes noncontextuality in the sense of Spekkens[106], whose classical case, embedding into a simplex, characterizes generalized-noncontextual models (Schmid, Selby, Wolfe, Kunjwal and Spekkens[107]). The present criterion

[91] Barnum and Wilce, *Local tomography*, note 15.

[92] van de Wetering, "Sequential product spaces are Jordan algebras", note 87; B. Westerbaan and J. van de Wetering, A computer scientist's reconstruction of quantum theory, *J. Phys. A* **55** (2022) 384002. doi:10.1088/1751-8121/ac8459. arXiv:2109.10707.

[93] L. Hardy, Quantum theory from five reasonable axioms, arXiv:quant-ph/0101012 (2001).

[94] Ll. Masanes and M. P. Müller, A derivation of quantum theory from physical requirements, *New J. Phys.* **13** (2011) 063001. arXiv:1004.1483.

[95] G. Chiribella, G. M. D'Ariano and P. Perinotti, Informational derivation of quantum theory, *Phys. Rev. A* **84** (2011) 012311. doi:10.1103/PhysRevA.84.012311. arXiv:1011.6451.

[96] R. D. Sorkin, Quantum mechanics as quantum measure theory, *Mod. Phys. Lett. A* **9** (1994) 3119. arXiv:gr-qc/9401003.

[97] C. Ududec, H. Barnum and J. Emerson, Three slit experiments and the structure of quantum theory, *Found. Phys.* **41** (2011) 396–405. doi:10.1007/s10701-010-9429-z. arXiv:0909.4787.

[98] Niestegge, *Conditional probability, three-slit experiments*, note 83.

[99] D. Finkelstein, J. M. Jauch, S. Schiminovich and D. Speiser, Foundations of quaternion quantum mechanics, *J. Math. Phys.* **3** (1962) 207–220. doi:10.1063/1.1703794.

[100] S. L. Adler, *Quaternionic Quantum Mechanics and Quantum Fields*, International Series of Monographs on Physics **88**, Oxford University Press, New York, 1995.

[101] A. Peres, Proposed test for complex versus quaternion quantum theory, *Phys. Rev. Lett.* **42** (1979) 683–686.

[102] M. McKague, M. Mosca and N. Gisin, Simulating quantum systems using real Hilbert spaces, *Phys. Rev. Lett.* **102** (2009) 020505. doi:10.1103/PhysRevLett.102.020505. arXiv:0810.1923.

[103] M.-O. Renou, D. Trillo, M. Weilenmann, T. P. Le, A. Tavakoli, N. Gisin, A. Acín and M. Navascués, Quantum theory based on real numbers can be experimentally falsified, *Nature* **600** (2021) 625–629. doi:10.1038/s41586-021-04160-4. arXiv:2101.10873.

[104] Müller and Garner, *Testing quantum theory*, note 71. Thm. 2.

[105] Garner and Müller, *Characterization of the probabilistic models*, note 71.

[106] R. W. Spekkens, Contextuality for preparations, transformations, and unsharp measurements, *Phys. Rev. A* **71** (2005) 052108. doi:10.1103/PhysRevA.71.052108. arXiv:quant-ph/0406166.

[107] D. Schmid, J. H. Selby, E. Wolfe, R. Kunjwal and R. W. Spekkens, Characterization of noncon-

fixes finitely many tests, compares all sequences under the Lüders update, and allows any dimension and rank; the criterion is speciality of the subalgebra the tests generate. No derivation of either statement from the other is established here.

## 9.2 Ideal measurements and the Lüders rule

The update $\rho \mapsto P\rho P$ for a projection of any rank is Lüders' rule[108], and Busch, Grabowski and Lahti[109] relate repeatability to ideality. Kleinmann[110] proves without complete positivity that a positive map sending the unit to $f$ and fixing every effect below $f$ exists only if $f$ is a projection $F$, and is then $X \mapsto FXF$. Chiribella and Yuan[111] define sharp measurements in any operational theory, and Cho[112] proves the Lüders form of ideal sharp instruments in comprehensive effectuses, which include von Neumann algebras.

For Jordan algebras, Niestegge identified Lüders conditioning with conditional probability and recovered the Jordan product from two complementary filters, $p \circ x = \frac{1}{2}(x + U_p x - U_{e-p} x)$[113], the identity (5.2) by which traces are read from words. His theorem[114] that $P(q \mid p) = r$ for all states if and only if $pqp = rp$ holds inside a given Hilbert space; the purity relations here run from abstract statistics to a representation.

Van de Wetering[115] showed that on a Euclidean Jordan algebra an invariant continuous sequential product is $U_{\sqrt{a}} b$. A Lean formalisation of sequential products[116] registers machine-checked classifications on $H_N(\mathbb{R})$ and $H_N(\mathbb{C})$, lists the octonionic classification as not formalised and proves Albert's power-associativity theorem for commutative Jordan rings carrying a real-module structure; a companion library[117] formalises the spectral theorem, the trace form and the Peirce decomposition of Euclidean Jordan algebras. Another Lean formalisation[118] proves the Jordan

---

textuality in the framework of generalized probabilistic theories, *PRX Quantum* **2** (2021) 010331. doi:10.1103/PRXQuantum.2.010331. arXiv:1911.10386.

[108] Lüders, "Über die Zustandsänderung durch den Meßprozeß", note 1.

[109] Busch, Grabowski and Lahti, *Repeatable measurements in quantum theory*, note 28.

[110] Kleinmann, *Sequences of projective measurements*, note 31. Sec. III A.

[111] G. Chiribella and X. Yuan, Measurement sharpness cuts nonlocality and contextuality in every physical theory, arXiv:1404.3348 (2014); Chiribella and Yuan, *Bridging the gap*, note 32.

[112] Cho, *Effectuses in Categorical Quantum Foundations*, note 33. Thm. 6.3.34.

[113] Niestegge, *Conditional probability, three-slit experiments*, note 83.

[114] G. Niestegge, Quantum probability's algebraic origin, *Entropy* **22** (2020) 1196. doi:10.3390/e22111196. arXiv:2009.08489.

[115] van de Wetering, "Three characterisations of the sequential product", note 36.

[116] B. Ehrlich, `sequential-products-lean`: classification of sequential products on Hermitian matrix algebras (Lean formalisation), repository `ehrlich-b/sequential-products-lean`, 2026; Palomar registry entries PALOMAR-2026-08-21-000009 and PALOMAR-2026-08-22-000006 (commit `9c6d92f5`), Apache License 2.0.

[117] B. Ehrlich, `euclidean-jordan`: Euclidean Jordan algebras in Lean 4, repository `ehrlich-b/euclidean-jordan`, 2026; Palomar registry entries PALOMAR-2026-08-23-000005, PALOMAR-2026-08-23-000006, PALOMAR-2026-08-29-000001 and PALOMAR-2026-08-29-000008, Apache License 2.0.

[118] M. J. Soares, The Albert algebra in Lean 4: Jordan identity and non-speciality, version 1.0.0, Zenodo, 22 July 2026, repository `markkasaurus/albert-algebra-lean`, Apache License 2.0. doi:10.5281/zenodo.21482444.

identity for $H_3(\mathbb{O})$, its real dimension 27, Glennie's identity in symmetrisations of associative algebras and an explicit violation of that identity in $H_3(\mathbb{O})$, from which non-speciality follows. The classifications of sequential products describe the update on a given algebra. The question here is which statistics of given ideal tests a Hilbert space can reproduce.

## 9.3 Temporal correlations and certificates

Temporal correlations entered as tests of macrorealism with the inequalities of Leggett and Garg[119], and Fritz[120] showed that the correlators of the temporal CHSH scenario with projective measurements on one system are those of the spatial scenario. The moment-matrix characterisation of Budroni, Moroder, Kleinmann and Gühne[121] is used in Section 3.4 and Lemma 6.4, and Vitagliano and Budroni[122] survey the method. Dual certificates for non-reproducible tables follow from this framework; Sections 4 and 6 use it for tables of $J$.

That a measured square with zero expectation kills a vector of the cyclic representation is standard in noncommutative optimisation (Pironio, Navascués and Acín[123]; Šupić and Bowles[124]), also for sequential measurements (Saha, Santos and Augusiak[125]). The present use differs in its output: a Jordan homomorphism of an abstractly given algebra, and Clifford relations among holonomies.

## 9.4 Isoclinic subspaces and algebras of projections

Wolf[126] put families of isoclinic planes into spheres spanned by anticommuting unitaries; Section 5.4 compares his Lemma 7[127] with the tetrahedron identity.

Popova and Samoilenko[128] study $*$-algebras generated by projections with $P_a P_b P_a = \tau_{ab} P_a$, over $\mathbb{C}$ and with orthogonality on non-edges, and decide existence of representations when the graph of nonorthogonal pairs is a tree; Samoilenko and Yershova[129] treat rank-one configurations. The four tests of Section 5.4 have all $\tau_{ab} > 0$ and satisfy three-letter relations as well.

[119] A. J. Leggett and A. Garg, Quantum mechanics versus macroscopic realism: is the flux there when nobody looks?, *Phys. Rev. Lett.* **54** (1985) 857–860. doi:10.1103/PhysRevLett.54.857.

[120] T. Fritz, Quantum correlations in the temporal Clauser–Horne–Shimony–Holt (CHSH) scenario, *New J. Phys.* **12** (2010) 083055. doi:10.1088/1367-2630/12/8/083055. arXiv:1005.3421.

[121] Budroni et al., "Bounding temporal quantum correlations", note 3.

[122] Vitagliano and Budroni, "Leggett–Garg macrorealism and temporal correlations", note 37.

[123] Pironio, Navascués and Acín, *Convergent relaxations*, note 42.

[124] I. Šupić and J. Bowles, Self-testing of quantum systems: a review, *Quantum* **4** (2020) 337. doi:10.22331/q-2020-09-30-337. arXiv:1904.10042.

[125] D. Saha, R. Santos and R. Augusiak, Sum-of-squares decompositions for a family of noncontextuality inequalities and self-testing of quantum devices, *Quantum* **4** (2020) 302. doi:10.22331/q-2020-08-03-302. arXiv:2002.12216.

[126] Wolf, "Geodesic spheres in Grassmann manifolds", note 11.

[127] Wolf, "Elliptic spaces in Grassmann manifolds", note 11.

[128] N. D. Popova and Yu. S. Samoilenko, On the existence of configurations of subspaces in a Hilbert space with fixed angles, *SIGMA* **2** (2006) 055. doi:10.3842/SIGMA.2006.055. arXiv:math/0605717.

[129] Yu. S. Samoilenko and Yu. Yu. Yershova, On equiangular configurations of subspaces of a Hilbert space, *Methods Funct. Anal. Topology* **12** (2006) 84–96.

## 9.5 Invariants of triangles and quadruples

As described by Ortega and Santander[130], Brehm[131] showed that a triangle in complex projective or hyperbolic space is determined up to isometry by its sides and a shape invariant, the real part of a product of three Hermitian products, and Hsiang[132] gave the trigonometry of all two-point homogeneous spaces of rank one, real, complex, quaternionic and octonionic.

Bargmann invariants of complex rank-one states give the Gram determinant of four pure states (Oszmaniec, Brod and Galvão[133]; see Section 5.4) and witness imaginarity (Fernandes, Wagner, Novo and Galvão[134]). Iltyakov[135] and Elduque and Iltyakov[136] study $F_4$-invariants of several elements of $J$ (compare Remark 5.4), and Schwarz[137] gives the invariant theory of $G_2$ and Spin(7).

## 9.6 Generation and speciality

Albert and Paige[138] showed that $J$ is not a homomorphic image of a special Jordan algebra.

Three primitive idempotents, as primitive axes of Jordan type $\frac{1}{2}$, generate an algebra of dimension at most nine (Gorshkov and Staroletov[139]); a JB-algebra generated by two projections and one further element, in particular by three projections, is special[140]; and four atoms can generate $J$[141],[142]. These results give speciality or generation; Proposition 5.1, Theorem 5.5 and Corollary 7.5 concern containment in a complex subalgebra, a quantitative criterion and the Helly number five. Earlier work gave a degree-nine scalar of three elements of $J$ whose nonvanishing implies that they generate $J$[143].

[130] Ortega and Santander, *Trigonometry of 'complex Hermitian' type*, note 10.

[131] Brehm, *The shape invariant of triangles*, note 44.

[132] W.-Y. Hsiang, On the laws of trigonometries of two-point homogeneous spaces, *Ann. Global Anal. Geom.* **7** (1989) 29–45. doi:10.1007/BF00137400.

[133] Oszmaniec, Brod and Galvão, *Measuring relational information*, note 53.

[134] C. Fernandes, R. Wagner, L. Novo and E. F. Galvão, Unitary-invariant witnesses of quantum imaginarity, *Phys. Rev. Lett.* **133** (2024) 190201. arXiv:2403.15066.

[135] A. V. Iltyakov, On generators of the algebra of polynomial invariants of group $F_4$, University of Sydney Research Report 97-4, 1997.

[136] A. Elduque and A. V. Iltyakov, On polynomial invariants of exceptional simple algebraic groups, *Canad. J. Math.* **51** (1999) 506–522. doi:10.4153/CJM-1999-023-2.

[137] G. W. Schwarz, Invariant theory of $G_2$ and $\mathrm{Spin}_7$, *Comment. Math. Helv.* **63** (1988) 624–663. doi:10.1007/BF02566782.

[138] Albert and Paige, "On a homomorphism property of certain Jordan algebras", note 62.

[139] I. Gorshkov and A. Staroletov, On primitive 3-generated axial algebras of Jordan type, *J. Algebra* **563** (2020) 74–99. doi:10.1016/j.jalgebra.2020.07.014. arXiv:2005.13791. Thm. 1.

[140] Roelands and Tiersma, "On the Shirshov–Cohn theorem for JB-algebras", note 9. Thm. 6.3.

[141] Ibid. Lemma 6.5.

[142] I. Gorshkov, A. Mamontov and A. Staroletov, On Jordan algebras that are factors of Matsuo algebras, *Int. Electron. J. Algebra* **37** (2025) 70–90. doi:10.24330/ieja.1488467. arXiv:2305.10958. Prop. 6.1.

[143] R. R. MacGregor, Finite observations and recognition of the compact Albert algebra, Palomar registry entry PALOMAR-2026-10-16-000013, 16 October 2026, Lean repository `macgregor-sudo/compact-albert-recognition`, commit `40d3b34`.

## 9.7 Instruments and memory

The dilation results of Davies and Lewis[144], Ozawa[145] and Khandelwal and Tavakoli[146] are compared with Proposition 8.6 in Section 8.2. Monràs and Winter[147] characterise finite completely positive realizations of finite-alphabet stationary processes using a semidefinite mapping cone compatible with their dynamics. A semidefinite description of states alone does not supply that compatibility. Fanizza, Lumbreras and Winter[148] give processes with finite generalised memory and no finite-dimensional quantum realization, and hence finitely correlated states in the sense of Fannes, Nachtergaele and Werner[149] that are not C*-finitely correlated; Theorem 8.1 is a finite-dimensional realization for the statistics of a theory that does not embed in quantum theory.

## 9.8 Octonionic quantum mechanics

Günaydin, Piron and Ruegg[150] built octonionic quantum mechanics on the Moufang plane, with points the atoms of $J$ and ideal measurements of the first kind for yes–no experiments. In their explicit configuration of Albert projections, the three intersections of corresponding sides of two perspective triangles fail to be collinear[151]. Catto expresses point–line incidence by $\mathrm{tr}(r \circ (p \times q)) = 0$, where $\times$ is the Freudenthal product[152], which goes back to Freudenthal's work on the octave plane[153]. For distinct atoms the trace-one representative of this line's polar point is $2(p \times q)/(1 - o_{pq})$. Günaydin, Piron and Ruegg also showed that two compatible experiments give order-independent results and identified the complex Jordan matrices as a Desarguesian subgeometry realised in a three-dimensional Hilbert space. The finite-word theorems here concern sequences of incompatible tests.

Finite sets of atoms of $J$, viewed as points of the octonionic projective plane, have also been studied as designs. Nasmith gives a common octonionic construction of the two strictly projective tight 5-designs, one of which consists of 819 points of that plane[154]. Höhn reconstructs the traceless Albert product from the third moment of

[144] Davies and Lewis, "An operational approach to quantum probability", note 74.

[145] Ozawa, "Quantum measuring processes of continuous observables", note 75.

[146] Khandelwal and Tavakoli, *Simulating quantum instruments*, note 76.

[147] Monràs and Winter, *Quantum learning of classical stochastic processes*, note 41. Thm. 21.

[148] M. Fanizza, J. Lumbreras and A. Winter, Quantum theory in finite dimension cannot explain every general process with finite memory, *Commun. Math. Phys.* **405** (2024) 50. doi:10.1007/s00220-023-04913-4. arXiv:2209.11225. (Journal Theorems 3–4 are Theorems 7–8 of arXiv v2.).

[149] M. Fannes, B. Nachtergaele and R. F. Werner, Finitely correlated states on quantum spin chains, *Commun. Math. Phys.* **144** (1992) 443–490. doi:10.1007/BF02099178.

[150] Günaydin, Piron and Ruegg, *Moufang plane and octonionic quantum mechanics*, note 29.

[151] Ibid. §4, pp. 78–80.

[152] S. Catto, Exceptional projective geometries and internal symmetries, arXiv:hep-th/0302079 (2003). eqs. (102), (115).

[153] H. Freudenthal, Beziehungen der $E_7$ und $E_8$ zur Oktavenebene I, *Indag. Math.* **16** (1954) 218–230 (= Proc. Kon. Ned. Akad. Wet. A **57**). doi:10.1016/S1385-7258(54)50032-6.

[154] B. Nasmith, Octonions and the two strictly projective tight 5-designs, *Algebr. Comb.* **5** (2022) 401–411. doi:10.5802/alco.215; B. Nasmith, Octonion integers and tight 5-designs, arXiv:2209.02398v1 (2022).

a projective 3-design, extends angle-preserving bijections between such designs to automorphisms in $F_4$, and so obtains the uniqueness of the 819-point design[155]; a second paper treats definite integral Albert algebras[156]. Those configurations are highly symmetric. The quadruples studied here are arbitrary, and none of these results is used.

The Albert algebra appears in proposals about internal degrees of freedom of particles, among them Günaydin and Gürsey[157], Dubois-Violette and Todorov[158], Krasnov[159], Boyle and Farnsworth[160], Boyle[161] and Baez and Schwahn[162]; Baez's surveys[163] explain the background. These works concern symmetry groups, not measurement statistics. Baez, Bokor and Boyle[164] treat pure states, expectation values and time evolution on Hermitian Jordan triples, with no rule for the state after a measurement. Exceptional groups acting by unitary representations on a Hilbert space are untouched by the present results, which constrain only theories that use $J$ as their structure of states and measurements. Pineda Montoya and Rua Munoz[165] represent quaternionic quantum mechanics exactly on a doubled complex space selected by an antiunitary symplectic structure, as the embedding $H_3(\mathbb{H}) \to H_6(\mathbb{C})$ does here, and compare candidate octonionic frameworks by the operational structure each preserves.

[155]G. Höhn, The unique extremal threemodular lattice of rank 26, the generalized hexagon (2,8), and the tight Cayley-plane 5-design, arXiv:2609.25086v1 (2026).

[156]G. Höhn, Definite integral Albert algebras: arithmetic and finite geometry, arXiv:2609.27550v1 (2026).

[157]M. Günaydin and F. Gürsey, Quark structure and octonions, *J. Math. Phys.* **14** (1973) 1651–1667. doi:10.1063/1.1666240.

[158]M. Dubois-Violette, Exceptional quantum geometry and particle physics, *Nucl. Phys. B* **912** (2016) 426–449. doi:10.1016/j.nuclphysb.2016.04.018. arXiv:1604.01247; M. Dubois-Violette and I. Todorov, Exceptional quantum geometry and particle physics II, *Nucl. Phys. B* **938** (2019) 751–761. doi:10.1016/j.nuclphysb.2018.12.012. arXiv:1808.08110.

[159]K. Krasnov, SO(9) characterization of the Standard Model gauge group, *J. Math. Phys.* **62** (2021) 021703. doi:10.1063/5.0039941. arXiv:1912.11282.

[160]L. Boyle and S. Farnsworth, The standard model, the Pati–Salam model, and "Jordan geometry", *New J. Phys.* **22** (2020) 073023. doi:10.1088/1367-2630/ab9709. arXiv:1910.11888.

[161]L. Boyle, The standard model, the exceptional Jordan algebra, and triality, *J. Math. Phys.* **67** (2026) 071701. doi:10.1063/5.0336175. arXiv:2006.16265.

[162]Baez and Schwahn, *The Standard Model gauge group*, note 5.

[163]Baez, "The octonions", note 5; J. C. Baez, Division algebras and quantum theory, *Found. Phys.* **42** (2012) 819–855. doi:10.1007/s10701-011-9566-z. arXiv:1101.5690.

[164]J. C. Baez, E. Bokor and L. Boyle, Jordan pair quantum theory and the standard model, arXiv:2607.10833 (2026).

[165]S. Pineda Montoya and J. H. Rua Munoz, Operational foundations for quaternionic and octonionic quantum models: exact quaternionic channels and error correction, para-linear operators, and categorical closure boundaries beyond associativity, arXiv:2608.20259 (2026).

# 10 Discussion

Each complex subalgebra on its own looks conventional (Section 5.5), and the exceptional structure is in how updates in different complex subalgebras compose. Could exceptional structure then act inside ordinary quantum theory as a compatibility rule across overlapping associative descriptions, with the seven-step gap as the cost of switching between them, rather than as a rival system?

### The box and its quantum realizations

For the box of Section 1.1 with the table of M or S, a completely positive realization with a 68-dimensional memory and one fixed completely positive instrument per button gives the same answer statistics for every finite sequence of presses, even when the next button is chosen from the earlier answers (Corollary 8.2).

For that family of tests and that preparation an exceptional description of the observable process coexists with a quantum implementation whose measurement structure is different. The Albert tests satisfy both (Id) and (Sh) on the closed convex hull of their reachable states, even when every affine effect on that state space is allowed, as shown in the companion paper (section "Ideality on reachable states"). This does not mean that this realization implements all those effects or encodes those states affinely, and it does not certify ideality on every state of the memory: Definition 3.2 requires non-disturbance on the full certainty faces, together with sharpness. For sharp effects, the states on which a certain answer disturbs later statistics lie outside the closed convex hull of the states the instruments themselves prepare (Remark 6.22).

Among ideal tests on Euclidean Jordan algebras, a table that violates a separating inequality, such as one on words of length seven, comes from tests on an algebra with an Albert summand of positive weight in the state (Remark 4.15). Outside that class a violation shows only that the tests are not ideal tests on a Hilbert space, and the realization of Corollary 8.2 is an exact quantum instance of instruments that are not ideal tests in the sense of Definition 3.2. Excluding the second requires ideality on whole certainty faces, not only on the states the preparations reach, and Proposition 6.20 states the calibration this requires.

### Open problems

1. *Optimal separation.* For four atoms that generate $J$ the minimum separating word length from $e/3$ is seven (Theorem 6.1, Proposition 6.2), but the uniform proof gives no explicit margin. What is the optimal separation at length seven, with a fixed normalisation of the witness $\Lambda$ of Proposition 6.19, and is it largest at the maximiser of $\mathcal{A}$ (Proposition 5.12)? Is there a margin bounded below by an explicit function of $\mathcal{A}$?

2. *Two schedules.* No single schedule of seven tests separates, even together with all words of length at most six, while two words of length seven from different schedules, with all words of length at most five, exclude every mixture of

quaternionic quadruples (Proposition 6.13). Do two schedules, together with all words of length at most six, suffice against every sequential projective representation?

3. *A separating linear equality.* Does some linear equality among word probabilities hold in every sequential projective representation and fail for $J$? None exists through length six from $e/3$ for any generating quadruple (Proposition 6.2), or through length five for any four atoms from any state (Remark 6.18).

4. *Mixtures of exceptional summands.* Can the primitivity hypothesis be removed from Theorem 4.11(b)? The simplest open case is a mixture of two copies of $J$ (Remark 4.12).

5. *Least dimension of a completely positive realization.* For four generating atoms the least dimension $d_{\min}$ of a completely positive realization from $e/3$ satisfies $6 \leq d_{\min} \leq 68$ (Proposition 8.4, Corollary 8.2). What is it, and is the spinor realization of Theorem 8.1 optimal in any sense? As more tests are admitted, with the preparation fixed and one completely positive instrument per test, do the least dimensions for exact matching grow without bound? A common bound would give one finite-dimensional completely positive realization for all tests at once, by the compactness argument of the companion paper (section "A fixed memory bound across menus"); no lower bound on the growth is proved here.

6. *Five tests.* Every four of the five tests of Theorem 7.2 are representable and the five are not, and five is the worst case (Corollary 7.5). This is a local-to-global failure of sequential representations, not of assignments of outcomes. Is there a description of it, comparable to the sheaf-theoretic description of contextuality, that explains the number five?

7. *The Lüders rule on one outcome, and approximate ideality.* Does the exclusion survive if the Lüders update is required only on the outcome "yes"? The tetrahedron argument uses the "no" outcomes through (5.2). Proposition 6.20 gives a stability version under approximate sharpness and uniform non-disturbance on both rounded spectral faces. Can these all-state calibration hypotheses be replaced by weaker, operationally testable conditions? A variant for generalised Lüders instruments of nearly projective effects is in the companion paper (section "Explicit margins for M and S").

# A Octonion and Jordan conventions; normal forms

This appendix fixes the conventions of the paper and proves the normal-form statements used in Section 5. Standard references are Baez[166], Springer and Veldkamp[167], McCrimmon[168] and Faraut and Korányi[169]; for the classification and representations of semisimple Jordan algebras, Jacobson and Jacobson[170], and for $F_4$ and the other exceptional groups, Yokota[171].

## A.1 Imported results

Tables 2 and 3 list the external theorems that the proofs use without proof. Results cited only for context (for instance Racine's classification of maximal subalgebras[172], which also implies Lemma E.1 for $J$), or proved here (for instance the existence of four generating atoms[173]), are not listed.

## A.2 Octonions

The octonions are built from the quaternions $\mathbb{H} = \operatorname{span}_{\mathbb{R}}\{1, i, j, ij\}$ by the Cayley–Dickson construction:

$$\mathbb{O} = \mathbb{H} \oplus \mathbb{H}\ell, \qquad (a + b\ell)(c + d\ell) = (ac - \bar{d}b) + (da + b\bar{c})\ell \qquad (a, b, c, d \in \mathbb{H}).$$

The basis is $e_0 = 1$, $e_1 = i$, $e_2 = j$, $e_3 = ij$, $e_4 = \ell$, $e_5 = i\ell$, $e_6 = j\ell$, $e_7 = (ij)\ell$. Conjugation is $\overline{a + b\ell} = \bar{a} - b\ell$, $|x|^2 = x\bar{x}$, and $\langle x, y\rangle = \operatorname{Re}(x\bar{y})$ is the Euclidean inner product of $\mathbb{R}^8$. The norm is multiplicative, so $\mathbb{O}$ has no zero divisors. Write $\mathbb{C}_i = \operatorname{span}\{1, i\}$ and $\operatorname{Im}\mathbb{O}$ for the imaginary octonions.

**Associators.** The associator $[x, y, z] = (xy)z - x(yz)$ is trilinear and alternating, and vanishes if one argument is real. By Artin's theorem (F5), $[x, y, z] = 0$ whenever $z$ lies in the subalgebra generated by $x$ and $y$. In the present basis $[i, j, \ell] = 2(ij)\ell$, so $|[i, j, \ell]|^2 = 4$. For imaginary $u, v$ put $u \times v = \operatorname{Im}(uv)$; then $uv = -\langle u, v\rangle + u \times v$ and $uv + vu = -2\langle u, v\rangle$, the Clifford relations of $\mathbb{R}^7$ with its negative definite form.

[166] Baez, "The octonions", note 5.

[167] T. A. Springer and F. D. Veldkamp, *Octonions, Jordan Algebras and Exceptional Groups*, Springer Monographs in Mathematics, Springer, Berlin, 2000. doi:10.1007/978-3-662-12622-6.

[168] K. McCrimmon, *A Taste of Jordan Algebras*, Universitext, Springer, New York, 2004.

[169] Faraut and Korányi, *Analysis on Symmetric Cones*, note 16.

[170] F. D. Jacobson and N. Jacobson, Classification and representation of semi-simple Jordan algebras, *Trans. Amer. Math. Soc.* **65** (1949) 141–169. doi:10.1090/S0002-9947-1949-0029367-8.

[171] I. Yokota, Exceptional Lie groups, arXiv:0902.0431 (2009).

[172] M. L. Racine, Maximal subalgebras of exceptional Jordan algebras, *J. Algebra* **46** (1977) 12–21. doi:10.1016/0021-8693(77)90391-X.

[173] Roelands and Tiersma, "On the Shirshov–Cohn theorem for JB-algebras", note 9; Gorshkov, Mamontov and Staroletov, *Jordan algebras that are factors of Matsuo algebras*, note 142.

Table 2: External theorems used in the proofs: Jordan algebras.

| Result | Source | Used in |
|---|---|---|
| Spectral theorem; closed self-dual cone; $U_a$ self-adjoint and positive; fundamental formula; Peirce decomposition and rules | [Faraut and Korányi, *Analysis on Symmetric Cones*, note 16, Thms. III.1.1, III.1.2, III.2.1, Props. II.3.3, III.2.2, IV.1.1, Thm. IV.2.1] | Secs. 3–8, Apps. B, D |
| Decomposition into simple ideals; classification of simple Euclidean Jordan algebras | [Jordan, von Neumann and Wigner, *On an algebraic generalization*, note 4]; [Faraut and Korányi, *Analysis on Symmetric Cones*, note 16, Prop. III.4.4, Thm. V.3.7] | Thms. 4.1, 4.11; Cor. 4.13; Lem. E.1; Cor. 7.5 |
| $J$ is not a homomorphic image of a special Jordan algebra | [Albert and Paige, "On a homomorphism property of certain Jordan algebras", note 62] | Prop. 6.17(c) |
| $[L_x, L_y]$ is a derivation of every Jordan algebra | standard; [McCrimmon, "The role of identities in Jordan algebras", note 60, §11] | Thm. F.6 |
| Transitivity of automorphisms on Jordan frames | [Faraut and Korányi, *Analysis on Symmetric Cones*, note 16, Thm. IV.2.5] | Lems. E.1, E.3 |
| Generic minimal equation of $J$; composition law $(U_a x)^\# = U_{a^\#} x^\#$ | [Springer and Veldkamp, *Octonions, Jordan Algebras and Exceptional Groups*, note 167, Ch. 5]; [McCrimmon, *A Taste of Jordan Algebras*, note 168, II.4.2.2(3)] | App. A; App. D |

**The Harvey–Lawson identity.** For $u, v, w \in \operatorname{Im}\mathbb{O}$,

$$\det\begin{pmatrix} \langle u,u\rangle & \langle u,v\rangle & \langle u,w\rangle \\ \langle v,u\rangle & \langle v,v\rangle & \langle v,w\rangle \\ \langle w,u\rangle & \langle w,v\rangle & \langle w,w\rangle \end{pmatrix} = \langle u\times v, w\rangle^2 + \tfrac{1}{4}\big|[u,v,w]\big|^2 \tag{A.1}$$

(Harvey and Lawson[174]; in this normalisation Salamon and Walpuski[175]).

## A.3 The Albert algebra

$J = H_3(\mathbb{O})$ is the space of Hermitian $3\times 3$ octonion matrices,

$$x = \begin{pmatrix} \xi_1 & x_3 & \bar{x}_2 \\ \bar{x}_3 & \xi_2 & x_1 \\ x_2 & \bar{x}_1 & \xi_3 \end{pmatrix}, \qquad \xi_k \in \mathbb{R},\ x_k \in \mathbb{O}, \tag{A.2}$$

so $x_1, x_2, x_3$ occupy the slots $23, 31, 12$. The Jordan product is $x\circ y = \frac{1}{2}(xy+yx)$ (matrix products, octonion entries), $e = \operatorname{diag}(1,1,1)$, $\operatorname{tr} x = \xi_1+\xi_2+\xi_3$, and

$$\langle x, y\rangle = \operatorname{tr}(x\circ y) = \textstyle\sum_k \xi_k\eta_k + 2\sum_k \langle x_k, y_k\rangle$$

is positive definite and associative. $E_{kk}$ are the diagonal matrix units, and $F_k(a)$ denotes the element with $a\in\mathbb{O}$ in slot $k$ (and $\bar{a}$ in the transposed position) and zeros elsewhere.

[174] Harvey and Lawson, "Calibrated geometries", note 46. Thm. IV.1.6.

[175] Salamon and Walpuski, "Notes on the octonions", Table 3. Lemma 4.4.

Table 3: External theorems used in the proofs: octonions, $F_4$, analysis and standard facts.

| Result | Source | Used in |
|---|---|---|
| Facts (F1)–(F3) on $F_4$; $F_4$ connected, stabiliser of an atom Spin(9), $\mathbb{O}P^2 = F_4/\mathrm{Spin}(9)$ | [Baez, "The octonions", note 5, §§3–4]; [Springer and Veldkamp, *Octonions, Jordan Algebras and Exceptional Groups*, note 167, Ch. 7]; [Harvey and Lawson, "Calibrated geometries", note 46, §IV.1] | App. A (normal forms); App. D (orientation) |
| Basic triples (F4); Artin's theorem and (F5); Cayley–Dickson doubling | [Baez, "The octonions", note 5, §4.1]; [Springer and Veldkamp, *Octonions, Jordan Algebras and Exceptional Groups*, note 167, Ch. 1] | App. A; App. C.4; Prop. 7.4 |
| Action of Spin(8) on the slots by vector and half-spin representations | [Baez, "The octonions", note 5, §§3.4, 4.1]; [A. Kollross, Octonions, triality, the exceptional Lie algebra $F_4$ and polar actions on the Cayley hyperbolic plane, *Internat. J. Math.* **31** (2020) 2050051. doi:10.1142/S0129167X20500512. arXiv:1802.08075] | Lem. E.3 |
| Harvey–Lawson identity for imaginary octonions | [Harvey and Lawson, "Calibrated geometries", note 46, Thm. IV.1.6]; [D. A. Salamon and T. Walpuski, Notes on the octonions, in *Proceedings of the 23rd Gökova Geometry–Topology Conference* (2017) 1–85. arXiv:1005.2820v4, Lemma 4.4] | Thm. 5.5 |
| Spectral theorem and trace-class duality over $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$ | standard | Thm. 3.3(b) |
| Tychonoff; strict separation of disjoint compact convex sets in $\mathbb{R}^N$; a continuous function on a compact set attains its minimum | standard | Cor. 4.16; Thm. 6.1(b) |
| Cartan–Dieudonné; Clifford representations of spin factors; the real Clifford algebra $\mathrm{Cl}_{0,3}$ and its conjugation; symplectic embedding $H_n(\mathbb{H}) \to H_{2n}(\mathbb{C})$ | standard | App. D; App. B.8; Prop. 6.11; Thm. 7.2(a); Lem. 6.9 |

**The Freudenthal adjoint and the cubic norm.** The adjoint goes back to Freudenthal's work on the octave plane[176].

$$N(x) = \xi_1\xi_2\xi_3 - \xi_1|x_1|^2 - \xi_2|x_2|^2 - \xi_3|x_3|^2 + 2\,\mathrm{Re}\big((x_1x_2)x_3\big),$$

and $x^\#$ has diagonal entries $\xi_2\xi_3 - |x_1|^2$, $\xi_3\xi_1 - |x_2|^2$, $\xi_1\xi_2 - |x_3|^2$ and slots

$$(x^\#)_1 = \overline{x_2x_3} - \xi_1x_1, \qquad (x^\#)_2 = \overline{x_3x_1} - \xi_2x_2, \qquad (x^\#)_3 = \overline{x_1x_2} - \xi_3x_3. \tag{A.3}$$

They satisfy $x \circ x^\# = N(x)e$ and the generic minimal equation $x^3 - \mathrm{tr}(x)x^2 + \mathrm{tr}(x^\#)x - N(x)e = 0$[177]. The eigenvalues of $x^\#$ are $\lambda_2\lambda_3, \lambda_3\lambda_1, \lambda_1\lambda_2$, so $x$ has rank at most one exactly when $x^\# = 0$.

**Lemma A.1** (Slots)**.**

*(a) For $a, b \in \mathbb{O}$: $F_1(a) \circ F_2(b) = \frac{1}{2}F_3(\overline{ab})$, $F_2(a) \circ F_3(b) = \frac{1}{2}F_1(\overline{ab})$, $F_3(a) \circ F_1(b) = \frac{1}{2}F_2(\overline{ab})$.*

*(b) For $\{k, l, m\} = \{1, 2, 3\}$ and $x \in J$: $U_{E_{kk}}x = \xi_kE_{kk}$, and $4E_{ll} \circ (E_{mm} \circ x)$ is the component of $x$ in the slot $lm$.*

*(c) Let $B \subseteq J$ be a unital subalgebra containing $E_{11}, E_{22}, E_{33}$ and $F_1(1), F_2(1), F_3(1)$, and let $S = \{a \in \mathbb{O} : F_1(a), F_2(a), F_3(a) \in B\}$. Then $S$ is a subalgebra of $\mathbb{O}$ closed under conjugation, and if $F_k(a) \in B$ for one $k$ then $a \in S$.*

*Narrower in Lean.* `jordanMul_slotElem01_slotElem12` and its analogues give (a) in the development's slot indexing; (b) and (c) are not stated in this form.

*Proof.* (a) In $F_1(a)F_2(b)$ the only nonzero entry is $ab$ at position $(2,1)$; in $F_2(b)F_1(a)$ it is $\bar{b}\bar{a} = \overline{ab}$ at $(1,2)$. The other two are the same computation with indices permuted cyclically. (b) $E_{mm}$ acts by 1 on $E_{mm}$, by $\frac{1}{2}$ on the two slots through $m$, and by 0 elsewhere; apply this twice.

(c) Take $F_k(a) \in B$, say $k = 3$. By (a) with the real elements $F_k(1)$, $F_1(\bar{a}) = 2F_2(1) \circ F_3(a)$ and $F_2(\bar{a}) = 2F_3(a) \circ F_1(1)$ lie in $B$, then $F_3(a) = 2F_1(\bar{a}) \circ F_2(1)$, $F_2(a) = 2F_3(1) \circ F_1(\bar{a})$ and $F_1(a) = 2F_2(\bar{a}) \circ F_3(1)$; the cases $k = 1, 2$ are the same. So $a, \bar{a} \in S$. If $a, b \in S$, then $F_3(\overline{ab}) = 2F_1(a) \circ F_2(b) \in B$, so $\overline{ab} \in S$ and hence $ab = \overline{\overline{ab}} \in S$. □

## A.4 Atoms and dyads

The atoms (Section 3.1) are the elements $p \geq 0$ with $\mathrm{tr}\,p = 1$ and $p^\# = 0$, that is, the elements with $p^2 = p$ and $\mathrm{tr}\,p = 1$. For an idempotent $c$ write $p \leq c$ for $p \in J_1(c)$. For $c = E_{11}$, $J_{1/2}$ is the sum of the slots 12 and 31 and $J_0$ is the 23 corner, a copy of $H_2(\mathbb{O})$. An idempotent of trace two is a sum of two orthogonal atoms.

[176] Freudenthal, "Beziehungen der $E_7$ und $E_8$ zur Oktavenebene I", note 153.

[177] Springer and Veldkamp, *Octonions, Jordan Algebras and Exceptional Groups*, note 167. Ch. 5.

**Dyads.** For $v = (v_1, v_2, v_3) \in \mathbb{O}^3$ put $(vv^\dagger)_{ab} = v_a \bar{v}_b$. If the entries of $v$ lie in an associative subalgebra and $\sum |v_a|^2 = 1$, then $vv^\dagger$ is an atom: the computation is that of a rank-one projection over an associative $*$-algebra. Conversely, let $p$ be an atom with $\xi_1 = p_{11} > 0$ and put

$$v = \left(\sqrt{\xi_1},\ \bar{x}_3/\sqrt{\xi_1},\ x_2/\sqrt{\xi_1}\right) \tag{A.4}$$

Then $p = vv^\dagger$: the entries in the first row and column agree by construction; the diagonal entries agree because $p^\# = 0$ gives $\xi_1\xi_2 = |x_3|^2$ and $\xi_1\xi_3 = |x_2|^2$; and the slot 23 agrees because $(p^\#)_1 = 0$ gives $x_1 = \overline{x_2 x_3}/\xi_1$. The entries of $v$ lie in the subalgebra generated by $x_2, x_3$, which is associative. In the same way $p^\# = 0$ gives: if $\xi_k = 0$, then the $k$-th row and column of $p$ vanish.

For dyads with entries in one associative subalgebra, $\langle vv^\dagger, ww^\dagger\rangle = |\sum_a \bar{v}_a w_a|^2$. For general octonionic dyads all traces are computed from the explicitly parenthesised matrix products, as in Lemma A.5.

## A.5 Imported facts about $F_4$ and $\mathbb{O}$

$F_4 = \mathrm{Aut}(J)$ is a compact connected Lie group of dimension 52; it preserves $e$, tr, $\langle\cdot,\cdot\rangle$, $N$ and $x \mapsto x^\#$. Two families of automorphisms are elementary: any automorphism of $\mathbb{O}$ applied entrywise (automorphisms of $\mathbb{O}$ commute with conjugation), and conjugation $x \mapsto \sigma x \sigma^{\mathsf{T}}$ by a real permutation matrix $\sigma$ (real entries associate with everything). The first fixes every real matrix, and the second permutes $E_{11}, E_{22}, E_{33}$. The following facts are used without proof (Baez[178]; Springer and Veldkamp[179]; Harvey and Lawson[180]).

(F1) $F_4$ is transitive on atoms. The stabiliser of $E_{11}$ acts on the traceless part of $J_0(E_{11})$ (the 23 corner with trace zero, $\cong \mathbb{R}^9$) through $SO(9)$.

(F2) The stabiliser of $E_{11}, E_{22}, E_{33}$ acts on each single slot through all of $SO(8)$.

(F3) Within that stabiliser, the subgroup fixing the real unit $F_3(1)$ is transitive on the unit sphere of each of the slots 23 and 31.

(F4) (Basic triples.) If $e_1, e_2, e_3$ are unit imaginary octonions with $e_2 \perp e_1$ and $e_3 \perp e_1, e_2, e_1e_2$, there is an automorphism of $\mathbb{O}$ with $i \mapsto e_1$, $j \mapsto e_2$, $\ell \mapsto e_3$.

(F5) The subalgebra generated by two octonions is associative (Artin) and contained in a quaternion subalgebra, i.e. a subalgebra isomorphic to $\mathbb{H}$.

From (F4): automorphisms of $\mathbb{O}$ act transitively on unit imaginary octonions; those fixing $i$ act transitively on unit vectors of $\mathbb{C}_i^\perp$; and there is one with $i \mapsto -i$, $j \mapsto j$, $\ell \mapsto \ell$, which acts on $\mathbb{C}_i$ by complex conjugation. The automorphisms fixing $\mathbb{H}$ pointwise act transitively on unit vectors of $\mathbb{H}\ell$: for a unit quaternion $q$, the map

[178] Baez, "The octonions", note 5. §§3–4.
[179] Springer and Veldkamp, *Octonions, Jordan Algebras and Exceptional Groups*, note 167. Chs. 1, 7.
[180] Harvey and Lawson, "Calibrated geometries", note 46. §IV.1.

$a + b\ell \mapsto a + (qb)\ell$ is an automorphism, as one checks from the multiplication rule ($\overline{qd}\, qb = \bar{d}b$ and $(qd)a + (qb)\bar{c} = q(da + b\bar{c})$).

A unital Jordan isomorphism between unital subalgebras of $J$ that contain a Jordan frame of $J$ preserves traces, because it maps frames to frames. If $\mathbb{H}' \subseteq \mathbb{O}$ is a quaternion subalgebra, then $H_3(\mathbb{H}')$ is a quaternionic subalgebra of $J$: an algebra isomorphism $\mathbb{H} \to \mathbb{H}'$ commutes with conjugation (in a unital subalgebra of $\mathbb{O}$, $\bar{x} = 2\operatorname{Re} x - x$, and $\operatorname{Re} x$ is read off the minimal equation $x^2 - 2\operatorname{Re}(x)\, x + |x|^2 = 0$) and preserves the norm, so applied entrywise it is a Jordan isomorphism $H_3(\mathbb{H}) \to H_3(\mathbb{H}')$, unital in $J$.

**Lemma A.2** (Restriction to a complex subalgebra)**.** *Let $W \subseteq J$ be a complex subalgebra, $\psi : W \to H_3(\mathbb{C})$ a unital Jordan isomorphism, $\rho$ a state of $J$ and $\rho_W$ its trace-orthogonal projection onto $W$. Then* $\operatorname{tr} x = \operatorname{Tr} \psi(x)$ *and* $\langle x, y\rangle = \operatorname{Tr}(\psi(x)\psi(y))$ *on $W$, $\psi$ maps the atoms of $J$ lying in $W$ to rank-one projections, $\psi(\rho_W)$ is a density matrix, and for every word $w = (q_1, \dots, q_m)$ in tests of atoms of $J$ lying in $W$*

$$\Pr_\rho(w) = \operatorname{Tr}\big(\psi(\rho_W)\, Q_1 \cdots Q_{m-1} Q_m Q_{m-1} \cdots Q_1\big), \qquad Q_t = \psi(q_t).$$

*In particular these statistics depend on $\rho$ only through $\rho_W$.*

*Formalized.* `complexRestriction Contract_holds`.

*Proof.* The diagonal idempotents of a Jordan frame of $W$ are three nonzero orthogonal idempotents of $J$ with sum $e$, hence atoms, since $\operatorname{tr} e = 3$. Every $x \in W$ is diagonalised by such a frame, so $\operatorname{tr} x = \operatorname{Tr} \psi(x)$, and $\langle x, y\rangle = \operatorname{Tr}(\psi(x)\psi(y))$ on $W$ because $\psi(x \circ y) = \frac{1}{2}(\psi(x)\psi(y) + \psi(y)\psi(x))$. In particular $\psi$ maps the atoms of $J$ lying in $W$ to rank-one projections. For $x \geq 0$ in $W$, $\langle \rho_W, x\rangle = \langle \rho, x\rangle \geq 0$, and the cone of $W$ is self-dual for the restricted trace form, so $\rho_W \geq 0$; also $\operatorname{tr} \rho_W = \langle \rho, e\rangle = 1$. The subalgebra $W$ is closed under $U_q$ for $q \in W$, and $\psi(U_a b) = \psi(a)\psi(b)\psi(a)$. So for every such word $w = (q_1, \dots, q_m)$, $\theta(w) \in W$ and

$$\Pr_\rho(w) = \langle \rho, \theta(w)\rangle = \langle \rho_W, \theta(w)\rangle = \operatorname{Tr}\big(\psi(\rho_W)\, Q_1 \cdots Q_{m-1} Q_m Q_{m-1} \cdots Q_1\big). \qquad \square$$

## A.6 Normal forms for three and four points

**Lemma A.3** (Three points)**.** *For any atoms $p_1, p_2, p_3$ of $J$ there is $g \in F_4$ with $gp_1 = E_{11}$, $gp_2 \in \operatorname{span}\{E_{11}, E_{22}\} + F_3(\mathbb{R}_{\geq 0})$, and $gp_3 \in H_3(\mathbb{C}_i)$ with real nonnegative slot-*31 *entry. The slot-*12 *entry of $gp_3$ may be taken with nonnegative coefficient of $-i$.*

*Different route.* `lemma_A_three`: automorphisms constructed in the development instead of the facts on $F_4$.

*Proof. Step 1.* By (F1) move $p_1$ to $E_{11}$.

*Step 2.* By (F1), rotate the traceless corner part of $p_2$ to a diagonal matrix, so $(p_2)_1 = 0$. Then $(p_2^\#)_1 = 0$ gives $\overline{x_2 x_3} = 0$, so $x_2 = 0$ or $x_3 = 0$; after the permutation of $E_{22}, E_{33}$ one may take $x_2 = 0$. If also $x_3 = 0$, $p_2$ is diagonal, hence one of $E_{11}, E_{22}, E_{33}$, and after the permutation $E_{11}$ or $E_{22}$. If $x_3 \neq 0$, the diagonal of $p_2^\#$ gives $\xi_1\xi_2 = |x_3|^2 > 0$ and $\xi_3\xi_1 = 0$, so $\xi_3 = 0$. Either way $p_2 \in \operatorname{span}\{E_{11}, E_{22}\} + F_3(\mathbb{O})$.

*Step 3.* By (F2), rotate the slot-12 entry of $p_2$ to a nonnegative real number. Let $K$ be the subgroup of (F3), fixing the frame and $F_3(1)$; it fixes $p_1$ and $p_2$.

*Step 4.* Write $p_3$ with entries $\xi_k, x_k$. Using $K$, rotate $x_2$ to a nonnegative real number. Then apply an automorphism of $\mathbb{O}$ entrywise, which fixes $p_1, p_2$ and the real $x_2$, to rotate $\operatorname{Im} x_1$ into $\mathbb{R}_{\geq 0} i$, so $x_1 \in \mathbb{C}_i$.

*Step 5.* If $\xi_3 \neq 0$, $(p_3^{\#})_3 = 0$ gives $x_3 = \overline{x_1 x_2}/\xi_3 \in \mathbb{C}_i$. If $\xi_3 = 0$, the third row and column of $p_3$ vanish, and an entrywise automorphism rotates $\operatorname{Im} x_3$ into $\mathbb{R}i$. In both cases $p_3 \in H_3(\mathbb{C}_i)$.

Finally, the entrywise automorphism with $i \mapsto -i$ (from (F4)) fixes $p_1, p_2$ and the real slot 31, keeps $p_3$ in $H_3(\mathbb{C}_i)$, and changes the sign of the $i$-component of its slot 12. □

**Lemma A.4** (Four points)**.** *Let $p_A, p_B, p_C, p_D$ be atoms with $0 < o_{AB} < 1$, $o_{AC} > 0$, $o_{AD} > 0$. There is $g \in F_4$ with $gp_i = v_i v_i^\dagger$ for unit vectors of the form* (5.6)*: $v_A = (1, 0, 0)$, $v_B = (B_0, B_1, 0)$, $v_C = (C_0, C_1 + C_2 i, C_3)$, $v_D = (D_0, X, Y)$ with $B_0, B_1, C_0, D_0 > 0$, $C_2, C_3 \geq 0$, $X = x_0 + x_1 i + x_2 j$, $Y = y_0 + y_1 i + y_2 j + y_3\, ij + y_4 \ell$, $x_2, y_4 \geq 0$. If no two of four atoms coincide or are orthogonal, the hypotheses hold.*

*Different route.* `lemma_A_normal`, `lemma_A_normal_nondegenerate`: automorphisms constructed in the development instead of the facts on $F_4$.

*Proof.* Apply Lemma A.3 to $p_A, p_B, p_C$ and take the vectors (A.4). For $p_B$: $B_0 = \sqrt{o_{AB}} > 0$, $B_1 = (p_B)_{12}/B_0 \geq 0$, and $B_1^2 = (p_B)_{22} = 1 - o_{AB} > 0$. For $p_C$: $C_0 = \sqrt{o_{AC}} > 0$, the second entry is the conjugate of the slot-12 entry over $C_0$, of the form $C_1 + C_2 i$ with $C_2 \geq 0$, and $C_3 = (p_C)_{31}/C_0 \geq 0$. For $p_D$: $D_0 = \sqrt{o_{AD}} > 0$ and $v_D = (D_0, X, Y)$ with arbitrary $X, Y \in \mathbb{O}$.

The automorphisms of $\mathbb{O}$ fixing $i$, applied entrywise, fix $p_A, p_B, p_C$, and one of them rotates the part of $X$ in $\mathbb{C}_i^\perp$ to a nonnegative multiple of $j$. Those fixing $\mathbb{H}$ pointwise fix $p_A, p_B, p_C$ and $X$, and one of them rotates the part of $Y$ in $\mathbb{H}\ell$ to a nonnegative multiple of $\ell$ (Section A.5). An entrywise automorphism $\alpha$ maps $vv^\dagger$ to $\alpha(v)\alpha(v)^\dagger$.

For the last sentence: if no pair coincides or is orthogonal, every $o_{ij} \in (0, 1)$, since $o_{ij} = 1$ iff $p_i = p_j$ and $o_{ij} = 0$ iff $p_i \circ p_j = 0$ (Lemma 3.1). □

**Lemma A.5** (The normal-form identity)**.** *Let $B_0, B_1, C_0, C_1, C_2, C_3, D_0 \in \mathbb{R}$ and $X, Y \in \mathbb{O}$ be arbitrary, let $v_A, \dots, v_D$ be the vectors* (5.6) *(not necessarily of unit length), $p_i = v_i v_i^\dagger$, and let $o, R, t$ be the thirteen traces of Definition 5.2 computed with the Jordan product of $J$. Put $z = C_1 + C_2 i$,*

*Narrower in Lean.* `sparseFourConfigOct_traces`, `sparseFourConfigOct_fourPointInvariant`. The trace identities and the formula for $\mathcal{A}$; the description of $G$ and $s$ is not stated.

$$g_1 = B_0 C_0 + B_1 z, \qquad g_2 = B_0 D_0 + B_1 X, \qquad g_3 = C_0 D_0 + \bar{z} X + C_3 Y,$$

*and $u_k = \operatorname{Im} g_k$. Then*

$$\begin{gathered} o_{AB} = B_0^2, \quad o_{AC} = C_0^2, \quad o_{AD} = D_0^2, \quad o_{BC} = |g_1|^2, \quad o_{BD} = |g_2|^2, \quad o_{CD} = |g_3|^2, \\ R_{ABC} = B_0 C_0 \operatorname{Re} g_1, \quad R_{ABD} = B_0 D_0 \operatorname{Re} g_2, \quad R_{ACD} = C_0 D_0 \operatorname{Re} g_3, \\ R_{BCD} = \operatorname{Re}\big((g_1 g_3)\bar{g}_2\big), \\ \omega_1 = C_0 D_0 \operatorname{Re}(g_1 \bar{g}_2), \quad \omega_2 = B_0 D_0 \operatorname{Re}(g_1 g_3), \quad \omega_3 = B_0 C_0 \operatorname{Re}(g_2 \bar{g}_3). \end{gathered} \tag{A.5}$$

*Where $B_0, C_0, D_0 > 0$, the matrix $G$ and the number $s$ of* (5.5) *are $G = (\langle u_k, u_l \rangle)$ and $s = \langle u_1 \times u_3, u_2 \rangle$. Finally*

$$\mathcal{A}(o, R, t) = \tfrac{1}{4}\, B_0^2 B_1^4\, C_0^2 C_2^2 C_3^2\, D_0^2 \left|[i, X, Y]\right|^2,$$

*an identity of polynomials over $\mathbb{Q}$ in 23 real variables.*

*Proof.* The proof first obtains the traces by sparse matrix contractions, then expands a scalar $3 \times 3$ determinant. No associative multiplication of octonion matrices is used. The scalar identities used below are conjugation reversal, $\mathrm{Re}(uv) = \mathrm{Re}(vu)$, $\mathrm{Re}((uv)w) = \mathrm{Re}(u(vw))$, and alternativity. These do not permit arbitrary reassociation of four octonions; the one such step needed below is justified explicitly.

*Transition probabilities and shape invariants.* Temporarily write $a = B_0$, $b = B_1$, $c = C_0$, $d = D_0$, $h = C_3$. The four columns are $(1, 0, 0)$, $(a, b, 0)$, $(c, z, h)$, $(d, X, Y)$, and

$$g_1 = ac + bz, \qquad g_2 = ad + bX, \qquad g_3 = cd + \bar{z}X + hY.$$

The pairings with $p_A = E_{11}$ are $a^2, c^2, d^2$. Since the column for $p_B$ is real, its pairings with $p_C, p_D$ are $|g_1|^2, |g_2|^2$. The remaining pairing is

$$\begin{aligned}\langle p_C, p_D \rangle = c^2d^2 + |z|^2|X|^2 + h^2|Y|^2 + 2cd\,\mathrm{Re}(\bar{z}X) + 2cdh\,\mathrm{Re}\,Y \\ + 2h\,\mathrm{Re}\big(z(Y\bar{X})\big) = |g_3|^2.\end{aligned}$$

The last cross term equals $2h\,\mathrm{Re}((\bar{z}X)\bar{Y})$ by conjugation and cyclicity of the real triple product. For dyads of columns $V, W$ with real first entries $v, w$, $(VV^\dagger \circ WW^\dagger)_{11} = vw\,\mathrm{Re}\sum_j \bar{V}_j W_j$. This proves the three triangle formulas involving $A$.

For the remaining triangle, write $B = (a, b, 0)$, $C = (c, z, h)$, $D = (d, X, Y)$ for columns. Because $B$ is real, contraction of the trace gives

$$R_{BCD} = \mathrm{Re}\sum_i \overline{(p_C B)_i}(p_D B)_i = \sum_i \mathrm{Re}\big((g_1\bar{C}_i)(D_i\bar{g}_2)\big).$$

Here $p_C B = C\bar{g}_1$ and $p_D B = D\bar{g}_2$. Real-part associativity first rewrites a summand as $\mathrm{Re}(((g_1\bar{C}_i)D_i)\bar{g}_2)$. For $i = 1, 3$, $C_i$ is real. For $i = 2$, both $g_1 = ac + bz$ and $\bar{C}_i = \bar{z}$ lie in the real span of $1, z$, so $[g_1, \bar{C}_i, D_i] = 0$ by alternativity. Thus the sum is $\mathrm{Re}((g_1 g_3)\bar{g}_2)$, as claimed.

*The double-pair traces.* For a column $V$ with real first entry $v$ and a Hermitian matrix $H$, the first-row rule is

$$\langle p_A \circ VV^\dagger, H \rangle = v^2 H_{11} + v\,\mathrm{Re}\sum_{j>1} V_j H_{1j}.$$

Multiplying just the indicated entries gives

$$(p_C \circ p_D)_{11} = cd\,\mathrm{Re}\,g_3, \qquad (p_C \circ p_D)_{12} = \tfrac{1}{2}\big(c(g_3\bar{X}) + d(\bar{g}_3\bar{z})\big).$$

The only nonscalar regroupings in the second formula are $(\bar{z}X)\bar{X} = \bar{z}(X\bar{X})$ and $(\bar{X}z)\bar{z} = \bar{X}(z\bar{z})$, both consequences of alternativity. The first-row rule therefore yields

$$\begin{aligned}2t_1 &= 2a^2cd\,\mathrm{Re}\,g_3 + abc\,\mathrm{Re}(g_3\bar{X}) + abd\,\mathrm{Re}(\bar{g}_3\bar{z}) \\ &= ac\,\mathrm{Re}(g_2\bar{g}_3) + ad\,\mathrm{Re}(g_1 g_3).\end{aligned}$$

For $j = 2, 3$, with $B_2 = b$ and $B_3 = 0$, the other required rows are

$$(p_B \circ p_D)_{1j} = \tfrac{1}{2}\big(a(g_2\bar{D}_j) + d\bar{g}_2 B_j\big), \qquad (p_B \circ p_C)_{1j} = \tfrac{1}{2}\big(a(g_1\bar{C}_j) + c\bar{g}_1 B_j\big).$$

Only real factors were moved here. Substitution gives

$$\begin{aligned} 2t_2 &= 2ac^2d\,\mathrm{Re}\,g_2 + ac\,\mathrm{Re}\big(z(g_2\bar{X}) + h(g_2\bar{Y})\big) + bcd\,\mathrm{Re}(z\bar{g}_2) \\ &= ac\,\mathrm{Re}(g_2\bar{g}_3) + cd\,\mathrm{Re}(g_1\bar{g}_2), \\ 2t_3 &= 2acd^2\,\mathrm{Re}\,g_1 + ad\,\mathrm{Re}\big(X(g_1\bar{z}) + hYg_1\big) + bcd\,\mathrm{Re}(X\bar{g}_1) \\ &= ad\,\mathrm{Re}(g_1g_3) + cd\,\mathrm{Re}(g_1\bar{g}_2). \end{aligned}$$

For the middle terms use real triple cyclicity to obtain, respectively, $\mathrm{Re}(g_2(\bar{g}_3 - cd))$ and $\mathrm{Re}(g_1(g_3 - cd))$. Taking $\omega_1 = t_2 + t_3 - t_1$, $\omega_2 = t_1 + t_3 - t_2$ and $\omega_3 = t_1 + t_2 - t_3$ proves the last three formulas of (A.5).

*The scalar determinant identity.* Now let $B_0, C_0, D_0 > 0$, and use the real scalars

$$\hat{x}_1 = \frac{R_{ABC}}{B_0C_0}, \quad \hat{x}_2 = \frac{R_{ABD}}{B_0D_0}, \quad \hat{x}_3 = \frac{R_{ACD}}{C_0D_0}, \qquad \hat{d}_1 = o_{BC}, \quad \hat{d}_2 = o_{BD}, \quad \hat{d}_3 = o_{CD},$$
$$\hat{\omega}_1 = \frac{\omega_1}{C_0D_0}, \quad \hat{\omega}_2 = \frac{\omega_2}{B_0D_0}, \quad \hat{\omega}_3 = \frac{\omega_3}{B_0C_0}, \qquad \hat{R} = R_{BCD}.$$

Then $G$ has diagonal $\hat{d}_1 - \hat{x}_1^2, \hat{d}_2 - \hat{x}_2^2, \hat{d}_3 - \hat{x}_3^2$ and off-diagonal entries $\hat{\omega}_1 - \hat{x}_1\hat{x}_2$, $\hat{x}_1\hat{x}_3 - \hat{\omega}_2$, $\hat{\omega}_3 - \hat{x}_2\hat{x}_3$. Put $\hat{L} = \hat{x}_3\hat{\omega}_1 + \hat{x}_2\hat{\omega}_2 + \hat{x}_1\hat{\omega}_3$, so that $s = \hat{R} + 2\hat{x}_1\hat{x}_2\hat{x}_3 - \hat{L}$. The elementary determinant formula reads

$$\begin{aligned} \det G = {} & (\hat{d}_1 - \hat{x}_1^2)(\hat{d}_2 - \hat{x}_2^2)(\hat{d}_3 - \hat{x}_3^2) + 2(\hat{\omega}_1 - \hat{x}_1\hat{x}_2)(\hat{x}_1\hat{x}_3 - \hat{\omega}_2)(\hat{\omega}_3 - \hat{x}_2\hat{x}_3) \\ & - (\hat{d}_1 - \hat{x}_1^2)(\hat{\omega}_3 - \hat{x}_2\hat{x}_3)^2 - (\hat{d}_2 - \hat{x}_2^2)(\hat{x}_1\hat{x}_3 - \hat{\omega}_2)^2 - (\hat{d}_3 - \hat{x}_3^2)(\hat{\omega}_1 - \hat{x}_1\hat{x}_2)^2. \end{aligned}$$

Group its terms by the number of $\hat{d}_k$ factors. Those with two or three give $\hat{d}_1\hat{d}_2\hat{d}_3 - \hat{x}_1^2\hat{d}_2\hat{d}_3 - \hat{x}_2^2\hat{d}_1\hat{d}_3 - \hat{x}_3^2\hat{d}_1\hat{d}_2$. Those with exactly one give

$$-\hat{d}_1\hat{\omega}_3^2 - \hat{d}_2\hat{\omega}_2^2 - \hat{d}_3\hat{\omega}_1^2 + 2\hat{d}_1\hat{\omega}_3\hat{x}_2\hat{x}_3 + 2\hat{d}_2\hat{\omega}_2\hat{x}_1\hat{x}_3 + 2\hat{d}_3\hat{\omega}_1\hat{x}_1\hat{x}_2;$$

the terms $\hat{d}_1\hat{x}_2^2\hat{x}_3^2, \hat{d}_2\hat{x}_1^2\hat{x}_3^2, \hat{d}_3\hat{x}_1^2\hat{x}_2^2$ cancel. The terms with no $\hat{d}_k$ give $-2\hat{\omega}_1\hat{\omega}_2\hat{\omega}_3 + (\hat{L} - 2\hat{x}_1\hat{x}_2\hat{x}_3)^2$. Indeed the off-diagonal product supplies the three cross terms of $\hat{L}^2$, $-2\hat{\omega}_1\hat{\omega}_2\hat{\omega}_3$, $-2\hat{x}_1\hat{x}_2\hat{x}_3\hat{L}$ and $2\hat{x}_1^2\hat{x}_2^2\hat{x}_3^2$; the remaining square terms supply $\hat{x}_3^2\hat{\omega}_1^2 + \hat{x}_2^2\hat{\omega}_2^2 + \hat{x}_1^2\hat{\omega}_3^2 - 2\hat{x}_1\hat{x}_2\hat{x}_3\hat{L} + 3\hat{x}_1^2\hat{x}_2^2\hat{x}_3^2$, and the diagonal product supplies $-\hat{x}_1^2\hat{x}_2^2\hat{x}_3^2$. Subtracting $s^2$ now gives

$$\begin{aligned} \det G - s^2 = {} & \hat{d}_1\hat{d}_2\hat{d}_3 - \hat{x}_1^2\hat{d}_2\hat{d}_3 - \hat{x}_2^2\hat{d}_1\hat{d}_3 - \hat{x}_3^2\hat{d}_1\hat{d}_2 - \hat{R}^2 - 4\hat{x}_1\hat{x}_2\hat{x}_3\hat{R} \\ & - \hat{d}_1\hat{\omega}_3^2 - \hat{d}_2\hat{\omega}_2^2 - \hat{d}_3\hat{\omega}_1^2 + 2\hat{\omega}_1(\hat{x}_1\hat{x}_2\hat{d}_3 + \hat{x}_3\hat{R}) + 2\hat{\omega}_2(\hat{x}_1\hat{x}_3\hat{d}_2 + \hat{x}_2\hat{R}) \\ & + 2\hat{\omega}_3(\hat{x}_2\hat{x}_3\hat{d}_1 + \hat{x}_1\hat{R}) - 2\hat{\omega}_1\hat{\omega}_2\hat{\omega}_3. \end{aligned}$$

Multiplication by $B_0^2C_0^2D_0^2$ gives the sixteen terms of (5.4). For example, the $\hat{x}_1^2\hat{d}_2\hat{d}_3$, $\hat{\omega}_1^2\hat{d}_3$ and $\hat{\omega}_1\hat{x}_1\hat{x}_2\hat{d}_3$ terms become $R_{ABC}^2o_{AD}o_{BD}o_{CD}$, $o_{AB}o_{CD}\omega_1^2$ and $\omega_1R_{ABC}R_{ABD}o_{CD}$, respectively. This proves the rational identity without using a relation among the thirteen traces.

*The associator.* The trace formulas give $G_{kl} = \langle u_k, u_l\rangle$ and $s = \langle u_1 \times u_3, u_2\rangle$ by separating real and imaginary parts. For example, $\mathrm{Re}(g_1\bar{g}_2) - \mathrm{Re}\,g_1\,\mathrm{Re}\,g_2 = \langle u_1, u_2\rangle$;

expanding the real triple product gives the stated expression for $s$. The Harvey–Lawson identity (A.1) therefore gives $\det G - s^2 = \frac{1}{4}|[g_1, g_3, g_2]|^2$. Finally, with the original octonion $z = C_1 + C_2 i$,

$$[g_1, g_3, g_2] = B_1^2\big([z, \bar{z}X, X] + C_3[z, Y, X]\big) = B_1^2 C_2 C_3 [i, Y, X].$$

The first associator vanishes in the associative subalgebra generated by $i, X$; real parts drop out of the second. This proves the last formula on the open set $B_0, C_0, D_0 > 0$. Both sides are real polynomials, so equality on this nonempty open set proves the identity for all the stated parameters. □

**Lemma A.6** (Degenerate pairs)**.** *If two of four atoms coincide or are orthogonal, the four atoms lie in a common quaternionic subalgebra of $J$.*

*Formalized.* `lemma_A_degenerate`.

*Proof. Coincident pair.* At most three distinct atoms remain; by Lemma A.3 they lie in $g^{-1}H_3(\mathbb{C}_i) \subset g^{-1}H_3(\mathbb{H})$.

*Orthogonal pair.* Call the pair $p_A, p_B$. By (F1) take $p_A = E_{11}$. Then $p_B \in J_0(E_{11})$ (Lemma 3.1), so $p_B = \frac{1}{2}(E_{22} + E_{33}) + y$ with $y$ traceless in the corner and $\langle y, y\rangle = \frac{1}{2}$; by (F1) rotate $y$ to $\frac{1}{2}(E_{22} - E_{33})$, so $p_B = E_{22}$. The frame stabiliser fixes both.

Now make $p_C$ real. If $(p_C)_{11} \neq 0$: by (F2) make its slot 12 real, by (F3) its slot 31, and then $(p_C^{\#})_1 = 0$ makes slot 23 real. If $(p_C)_{11} = 0 \neq (p_C)_{22}$: the first row and column vanish, and (F2) makes slot 23 real. If both vanish, $p_C = E_{33}$.

Finally, $p_D = vv^\dagger$ for the analogue of (A.4) built from a nonzero diagonal entry of $p_D$; this $v$ has one real entry and at most two octonion entries. By (F5) they lie in a quaternion subalgebra $\mathbb{H}'$, so $p_A, \ldots, p_D$ lie in the quaternionic subalgebra $H_3(\mathbb{H}')$ (Section A.5). □

**Lemma A.7** (The zero set)**.**

*(a) Any four atoms of the standard $H_3(\mathbb{H}) \subset J$ are mapped by a trace-preserving Jordan automorphism of $H_3(\mathbb{H})$ to dyads of vectors of the shape* (5.6) *with $B_0, B_1, C_0, C_3, D_0 \geq 0$ real, $C_1, C_2$ real and $X, Y \in \mathbb{H}$.*

*(b) In the situation of Lemma A.4, if $C_2 C_3 x_2 y_4 = 0$, the four atoms lie in a common quaternionic subalgebra.*

*Narrower in Lean.* (b) `lemma_A_zero_of_factor`. In place of (a), Lean proves its consequence that $\mathcal{A}$ vanishes on four atoms of a quaternionic subalgebra (`lemma_A_zero_of_quat`).

*Proof.* (a) Unitaries $U$ of $\mathbb{H}^3$ act on $H_3(\mathbb{H})$ by $x \mapsto UxU^*$, a trace-preserving Jordan automorphism, and a dyad $ww^*$ does not change under $w \mapsto wq$ for a unit quaternion $q$. Let $w_A, \ldots, w_D$ be unit vectors. A unitary maps $w_A$ to $(1, 0, 0)$. After a right phase the first entry of $w_B$ is real and nonnegative, and a unitary of the last two coordinates makes $w_B = (B_0, B_1, 0)$. A right phase makes the first entry of $w_C$ real and nonnegative, and $\mathrm{diag}(1, 1, r)$ with $|r| = 1$ its third entry. The unitary $qI$ acts by $x \mapsto qx\bar{q}$ entrywise, fixes real entries and rotates the imaginary part of the second entry of $w_C$ into $\mathbb{R}i$. A right phase makes the first entry of $w_D$ real and nonnegative.

(b) If $x_2 = 0$, then $X \in \mathbb{C}_i$ and all entries lie in the subalgebra generated by $i$ and $Y$; if $y_4 = 0$, they lie in $\mathbb{H}$; if $C_2 = 0$, the first three vectors are real and the entries lie in the subalgebra generated by $X$ and $Y$. In each case (F5) gives a quaternion subalgebra $\mathbb{H}'$ with all four atoms in $H_3(\mathbb{H}')$.

Let $C_3 = 0$. Then $p_A, p_B, p_C \le c := E_{11} + E_{22}$ in the normal form. Apply Lemma A.4 to the labelling $(A, B, D, C)$, giving $g'$; then $g'p_A = E_{11}$, $g'p_B = v'_B v'^{\dagger}_B$ with $v'_B = (B'_0, B'_1, 0)$, $B'_1 \neq 0$, and $g'p_D \in H_3(\mathbb{C}_i)$.

The idempotent $c' = g'c$ has trace two, and $E_{11}, g'p_B \le c'$. The claim is that $c' = E_{11} + E_{22}$. The element $q = c' - E_{11}$ is an idempotent of trace one orthogonal to $E_{11}$, hence an atom in $J_0(E_{11})$, so $q_{11} = 0$ and $q_{12} = 0$. Since $g'p_B \in J_1(c') \perp J_0(c') \ni e - c'$,

$$\langle g'p_B, q\rangle = \langle g'p_B, e - E_{11}\rangle = B'^2_1, \qquad \text{while} \qquad \langle g'p_B, q\rangle = B'^2_1\, q_{22}.$$

So $q_{22} = 1$, hence $q_{33} = 0$ and (by $q^{\#} = 0$) $q = E_{22}$.

Therefore $g'p_C \le E_{11} + E_{22}$, and its dyad vector (A.4) has the form $(C'_0, X', 0)$. All entries of the four new vectors lie in the subalgebra generated by $i$ and $X'$, and (F5) concludes. □

**Lemma A.8** (Generation)**.** *In the situation of Lemma A.4:*

*Formalized.* `lemma_A_generate_three`, `lemma_A_generate_four`.

*(a) if $C_2C_3 \neq 0$, then $p_A, p_B, p_C$ generate $H_3(\mathbb{C}_i)$;*

*(b) if moreover $x_2y_4 \neq 0$, then $p_A, \dots, p_D$ generate $J$.*

*Proof.* Let $B$ be the generated unital subalgebra. For an idempotent $c \in B$ the Peirce projections $U_c$, $U_{e-c}$ and $4L_c(\mathrm{id} - L_c)$ are polynomials in $L_c$, so $B$ contains the Peirce components of its elements with respect to $c$; once $E_{11}, E_{22}, E_{33} \in B$, it contains by Lemma A.1(b) their diagonal and slot components.

(a) $B \ni E_{11}$, and $p_B$ has diagonal $(B_0^2, B_1^2, 0)$ and slot 12 equal to $B_0B_1$. Its Peirce components with respect to $E_{11}$ are $B_0^2E_{11}$, $F_3(B_0B_1)$ and $U_{e-E_{11}}p_B = B_1^2E_{22}$. So $E_{22} \in B$, $E_{33} = e - E_{11} - E_{22} \in B$ and $F_3(1) \in B$. The slots of $p_C$ give $F_2(C_0C_3)$, hence $F_2(1)$, and $F_3(C_0(C_1 - C_2i))$, hence $F_3(i)$. Then $F_1(1) = 2F_2(1) \circ F_3(1)$, and Lemma A.1(c) gives $i \in S$, so $\mathbb{C}_i \subseteq S$ and $H_3(\mathbb{C}_i) \subseteq B$. Equality holds as all three atoms lie in $H_3(\mathbb{C}_i)$.

(b) The slot 12 of $p_D$ is $D_0\bar{X}$; subtracting its $\mathbb{C}_i$-part gives $F_3(x_2j) \in B$, so $j \in S$ and $\mathbb{H} \subseteq S$. The slot 31 of $p_D$ is $D_0Y$; subtracting its $\mathbb{H}$-part gives $F_2(D_0y_4\ell) \in B$, so $\ell \in S$. Since $i, j, \ell$ generate $\mathbb{O}$, $S = \mathbb{O}$ and $B = J$. □

**Lemma A.9** (Triangles in the normal form)**.** *Let $p_1, p_2, p_3, p_4$ be atoms of $J$ and $\Delta_W = \Delta_W(p_1, p_2, p_3)$ as in Section 5.*

*Narrower in Lean.* (b) `lemma_A_window_generated`; (c) `prop_windowdet_zero`. Part (a) and the identity $P_{gW}g = gP_W$ are not stated as separate theorems.

*(a) If no two of $p_1, p_2, p_3$ coincide or are orthogonal and $p_4$ is not orthogonal to all of them, then after relabelling $p_1, p_2, p_3$ so that $\langle p_1, p_4\rangle > 0$, Lemma A.4 applies to $(p_1, p_2, p_3, p_4)$ in the roles of $(A, B, C, D)$, and in that normal form $\Delta_W = B_0^2B_1^4C_0^2C_2^2C_3^2$.*

*(b) If $\Delta_W \neq 0$, the hypotheses of (a) hold, so $0 < o_{ij} < 1$ for $i, j \le 3$; moreover $W = \mathrm{Jord}\langle e, p_1, p_2, p_3\rangle$ satisfies $gW = H_3(\mathbb{C}_i)$ for the element $g \in F_4$ of Lemma A.4 in (a), so $W$ is a complex subalgebra, and $P_{gW}g = gP_W$.*

*(c) If $\Delta_W = 0$, then $\mathcal{A}(p_1, p_2, p_3, p_4) = 0$.*

*Proof.* (b), first part. If two of the three atoms coincide, $N(p_1+p_2+p_3)=0$ (with $p_1=p_2$ and $o=o_{13}$: $1-1-2o+2o=0$); if two are orthogonal, their Jordan product vanishes (Lemma 3.1), so $R_{123}=0$ and the second factor vanishes. So $\Delta_W\neq 0$ forces $0<o_{ij}<1$. Moreover $p_1+p_2+p_3\geq 0$ has $N\neq 0$, hence is invertible, so $p_4$ is not orthogonal to all three.

(a) $\Delta_W$, $W$ and $\mathcal{A}$ are symmetric in $p_1,p_2,p_3$, so it may be assumed that $\langle p_1,p_4\rangle>0$. Lemma A.4 applies to $(p_1,p_2,p_3,p_4)$ in the roles of $(A,B,C,D)$. The first three vectors have entries in $\mathbb{C}_i$, so the traces of the triple are those of complex unit vectors: $N(p_1+p_2+p_3)=|\det(v_1,v_2,v_3)|^2=B_1^2C_3^2$, and $o_{12}o_{23}o_{31}-R_{123}^2$ is the squared imaginary part of $\langle v_1|v_2\rangle\langle v_2|v_3\rangle\langle v_3|v_1\rangle=B_0C_0(B_0C_0+B_1C_1+B_1C_2i)$. Hence

$$\Delta_W=B_0^2B_1^4\,C_0^2C_2^2C_3^2.$$

(b), second part. As $\Delta_W\neq 0$, the formula of (a) gives $C_2C_3\neq 0$, and Lemma A.8(a) shows that $gp_1,gp_2,gp_3$ generate $H_3(\mathbb{C}_i)$, so $gW=H_3(\mathbb{C}_i)$. The automorphism $g$ is orthogonal, so $P_{gW}g=gP_W$, and $g$ preserves $N$.

(c) If two of $p_1,p_2,p_3$ coincide or are orthogonal, or $p_4$ is orthogonal to all three, Lemma A.6 and Theorem 5.5(ii) give $\mathcal{A}=0$. Otherwise relabel so that $\langle p_1,p_4\rangle>0$; then the formula of (a) holds in the normal form, so $C_2C_3=0$ and $\mathcal{A}=0$ by Lemma A.5. □

## A.7 The trace relation

In the identity $24e_4=s_1^4-6s_1^2s_2+3s_2^2+8s_1s_3-6s_4=0$ of the proof of Proposition 5.3, with $x=\sum_i\lambda_ip_i$ and $s_k=\operatorname{tr}(x^k)$, compare the coefficients of $\lambda_A\lambda_B\lambda_C\lambda_D$; terms containing a square $\lambda_i^2$ do not contribute. Since $\operatorname{tr}p_i=1$, $s_1=\sum\lambda_i$ and $s_2=\sum\lambda_i^2+2\sum_{i<j}\lambda_i\lambda_jo_{ij}$. In $s_3=\operatorname{tr}(x\circ x^2)$ the coefficient of $\lambda_i\lambda_j\lambda_k$, for distinct labels, is $6R_{ijk}$ by associativity of the trace form. In $s_4=\langle x^2,x^2\rangle$ only the cross terms $2\lambda_i\lambda_jp_i\circ p_j$ of $x^2$ contribute, each pairing of $\{A,B,C,D\}$ twice, giving $8(t_1+t_2+t_3)$. The coefficients of the other products are $24$ $(s_1^4)$, $4\sum_{i<j}o_{ij}$ $(s_1^2s_2)$, $8(o_{AB}o_{CD}+o_{AC}o_{BD}+o_{AD}o_{BC})$ $(s_2^2)$ and $6\sum R$ $(s_1s_3)$. So $0=24-24\sum o_{ij}+24\sum_\pi o\,o+48\sum R-48\sum t$, which is (5.3).

**There is no second polynomial relation.** A local inverse gives a direct proof of Remark 5.4(a). Use the strict part of the normal form, writing

$$v_B=(a,b,0),\quad v_C=(c,z,h),\quad v_D=(d,X,Y),$$
$$z=C_1+C_2i,\quad X=x_0+x_1i+x_2j,\quad Y=y_0+y_1i+y_2j+y_3ij+y_4\ell,$$

with unit columns and $a,b,c,C_2,h,d,x_2,y_4>0$. Take as independent input the six transition probabilities, $R_{ABC},R_{ABD},R_{ACD}$ and $\omega_1,\omega_2,\omega_3$. The trace formulas of Lemma A.5 recover the normal-form coordinates in the following order:

$$a=\sqrt{o_{AB}},\qquad b=\sqrt{1-a^2},\qquad c=\sqrt{o_{AC}},$$
$$C_1=\frac{R_{ABC}/(ac)-ac}{b},\quad C_2=\frac{\sqrt{o_{BC}-(ac+bC_1)^2}}{b},\quad h=\sqrt{1-c^2-C_1^2-C_2^2},$$
$$d=\sqrt{o_{AD}},\qquad x_0=\frac{R_{ABD}/(ad)-ad}{b},\qquad |X|^2=\frac{o_{BD}-(ad)^2-2adbx_0}{b^2}.$$

Put $u = ac + bC_1$, $v = bC_2$ and $q_0 = ad + bx_0$. Since $\omega_1 = cd(uq_0 + vbx_1)$, it determines $x_1$; then $x_2 = \sqrt{|X|^2 - x_0^2 - x_1^2} > 0$. Write $g_3 = m + ni + pj + qij + r\ell$. The next three equations determine $m, n, p$:

$$m = \frac{R_{ACD}}{cd}, \qquad \omega_2 = ad(um - vn), \qquad \omega_3 = ac(q_0 m + bx_1 n + bx_2 p).$$

Set $K = cd + \bar{z}X$. Its components in $1, i, j, ij$ are

$$K_0 = cd + C_1x_0 + C_2x_1, \quad K_1 = C_1x_1 - C_2x_0, \quad K_2 = C_1x_2, \quad K_3 = -C_2x_2.$$

The identity $g_3 = K + hY$ now determines $y_0, y_1, y_2$. Finally,

$$o_{CD} = |K|^2 + h^2(1 - d^2 - |X|^2) + 2h(K_0y_0 + K_1y_1 + K_2y_2 - C_2x_2y_3)$$

determines $y_3$, because its coefficient is nonzero, and normalization gives

$$y_4 = \sqrt{1 - d^2 - |X|^2 - y_0^2 - y_1^2 - y_2^2 - y_3^2} > 0.$$

Every division and square root is valid near the trace coordinates of the configuration S in (F.8): there $a = b = 1/\sqrt{2}$ and $c = C_2 = h = d = x_2 = y_4 = 1/\sqrt{3}$. The formulas are continuous there and realize every sufficiently nearby choice of the twelve input traces. Thus their image contains an open subset of $\mathbb{R}^{12}$.

The change between the three $t_r$ and the three $\omega_r$ is invertible. Equation (5.3), with sign chosen appropriately, is monic linear in $R_{BCD}$. Divide any polynomial relation among the thirteen traces by this polynomial. Its remainder is a polynomial in the other twelve coordinates and vanishes on the open set just constructed, so it is the zero polynomial. Hence every relation is a multiple of (5.3), and the trace image has dimension twelve. This proves the asserted principal-ideal statement without an irreducibility or rank calculation.

## A.8 The Clifford volume element

**Lemma A.10** (Volume element)**.** *Let $N_1, N_2, N_3$ be elements of a unital associative real algebra, with scalars interpreted as multiples of its unit, and with $N_aN_b + N_bN_a = -2M_{ab}$ for a real symmetric $3 \times 3$ matrix $M$, and put $\Gamma = \frac{1}{2}(N_1N_2N_3 - N_3N_2N_1)$. Then*

$$\Gamma = N_1N_2N_3 + M_{23}N_1 - M_{13}N_2 + M_{12}N_3, \qquad \Gamma^2 = \det M.$$

*Formalized.* `localUnit_volume_eq_half_reverse`, `localUnit_volume_square`, in an algebra with a local unit, which contains the unital case.

*Proof.* Moving factors with the anticommutators,

$$N_3N_2N_1 = -N_1N_2N_3 - 2M_{23}N_1 + 2M_{13}N_2 - 2M_{12}N_3,$$

which gives the first formula.

For the second, the relations reduce every word in the $N_a$ to a combination of the eight words $1, N_a, N_aN_b$ $(a < b), N_1N_2N_3$, with coefficients that are polynomials in the $M_{ab}$ determined by the reordering rule alone. So $\Gamma^2 - \det M = \sum_w c_w(M)\, w$ with polynomials $c_w$, the same in every associative algebra. It suffices to show

$c_w = 0$ for positive definite $M$, since a polynomial vanishing on an open set vanishes identically.

Let $M = LL^{\mathsf{T}}$ with $L$ invertible, and let $\gamma_1, \gamma_2, \gamma_3$ generate the real Clifford algebra with $\gamma_k^2 = -1$ and $\gamma_k\gamma_l = -\gamma_l\gamma_k$ ($k \neq l$), which has dimension 8 with basis the ordered $\gamma$-words. The elements $N_a = \sum_k L_{ak}\gamma_k$ satisfy the relations, and the eight words above in these $N_a$ are a basis (they are obtained from the $\gamma$-words by an invertible change of generators).

In $N_1N_2N_3 - N_3N_2N_1$, a word $\gamma_r\gamma_s\gamma_t$ with a repeated index cancels against its reversal, while for distinct indices reversal changes the sign. So $\Gamma = \det(L)\,\gamma_1\gamma_2\gamma_3$, and $(\gamma_1\gamma_2\gamma_3)^2 = 1$ gives $\Gamma^2 = \det M$. By linear independence, $c_w(M) = 0$. □

## A.9 The four-point invariant in full

Expanded in the thirteen traces, the polynomial (5.4) has 52 monomials: one sextic in the transition probabilities, four of type $R^2ooo$, one quartic in the $R$'s, eighteen of type $oo\,tt$, eighteen of type $RRo\,t$, and ten cubic in the $t$'s. To see this, substitute $\omega_r = \sigma - 2t_r$ with $\sigma = t_1 + t_2 + t_3$, so that for instance $\omega_1 = -t_1 + t_2 + t_3$. Each $\omega_r^2$ has all six quadratic monomials in the $t$'s with nonzero coefficients ($\pm1$, $\pm2$), and the three $\omega_r^2$ carry the three different factors $o_{AB}o_{CD}$, $o_{AC}o_{BD}$, $o_{AD}o_{BC}$: eighteen monomials. Each $\omega_r$ has three monomials, and the six products $RRo$ multiplying them are distinct: eighteen more. Finally $\omega_1\omega_2\omega_3 = -\sum_r t_r^3 + \sum_{r\neq r'} t_r^2 t_{r'} - 2t_1t_2t_3$ has all ten cubic monomials. Monomials of different groups differ in their $o$ and $R$ factors, so nothing cancels, and $1 + 4 + 1 + 18 + 18 + 10 = 52$.

**The compact form.** With $a, b, c$ as in (5.5) (where $o_{AB}o_{AC}o_{AD} > 0$) and $x_1 = R_{ABC}/ab$, $x_2 = R_{ABD}/ac$, $x_3 = R_{ACD}/bc$, the matrix and number in (5.5) are

$$G = \begin{pmatrix} o_{BC} - x_1^2 & \omega_1/bc - x_1x_2 & x_1x_3 - \omega_2/ac \\ \omega_1/bc - x_1x_2 & o_{BD} - x_2^2 & \omega_3/ab - x_2x_3 \\ x_1x_3 - \omega_2/ac & \omega_3/ab - x_2x_3 & o_{CD} - x_3^2 \end{pmatrix}, \qquad s = R_{BCD} - x_1x_2x_3 + x_2G_{13} - x_1G_{23} - x_3G_{12}. \tag{A.6}$$

The 24 relabellings of $A, B, C, D$ permute the terms of (5.4), since a relabelling permutes the pairings and hence the $\omega_r$. $F_4$-invariance holds because the traces are invariant.

# B Proofs for Sections 3 and 4

## B.1 Theorem 3.3(b): ideal tests on a Hilbert space

Let $\mathcal{S}_1$ be the real space of self-adjoint trace-class operators on $\mathcal{H}$ and $(T_1, T_0)$ positive real-linear maps on $\mathcal{S}_1$ with $\operatorname{Tr}(T_1X + T_0X) = \operatorname{Tr} X$. For $X \geq 0$, $0 \leq \operatorname{Tr} T_1X \leq \operatorname{Tr} X$, so $X \mapsto \operatorname{Tr} T_1X$ is bounded by the trace norm and, by the duality between self-adjoint trace-class and bounded self-adjoint operators, there is a bounded self-adjoint $A = T_1^*(1)$ with $\operatorname{Tr}(AX) = \operatorname{Tr}(T_1X)$ and $0 \leq A \leq 1$. (Over $\mathbb{R}$ and $\mathbb{H}$ as over $\mathbb{C}$: the bounded symmetric real bilinear form $b(\xi, \eta) = \frac{1}{2} \operatorname{Tr} T_1(|\xi\rangle\langle\eta| + |\eta\rangle\langle\xi|)$ is invariant under $\xi \mapsto \xi u$, $\eta \mapsto \eta u$ for unit scalars $u$, hence is $\operatorname{Re}\langle\xi, A\eta\rangle$ for a bounded self-adjoint $\mathbb{K}$-linear $A$; finite-rank operators are trace-norm dense.)

*Step 1: $A$ is a projection.* Let $P, Q, R$ be the spectral projections of $A$ for $\{1\}$, $\{0\}$ and $(0, 1)$. For a density operator $\varrho$, $\operatorname{Tr}(\varrho(1 - A)) = \|(1 - A)^{1/2}\varrho^{1/2}\|_2^2$ vanishes iff $\operatorname{ran}\varrho \subseteq \ker(1 - A) = \operatorname{ran} P$. So $F_1$ is the set of density operators supported in $P$, and $F_0$ those supported in $Q$. If $R \neq 0$, a unit vector $\xi \in \operatorname{ran} R$ gives the state $|\xi\rangle\langle\xi|$ and the effect $R$, which violate (Sh). So $R = 0$ and $A = P$.

*Step 2: diagonal blocks.* Write $P^\perp = 1 - P$. By (Id), $T_1$ fixes the density operators supported in $P$, which span $\{X \in \mathcal{S}_1 : X = PXP\}$. If $Y \geq 0$ and $Y = P^\perp YP^\perp$, then $T_1Y \geq 0$ and $\operatorname{Tr} T_1Y = \operatorname{Tr}(PY) = 0$, so $T_1Y = 0$; such $Y$ span $\{X : X = P^\perp XP^\perp\}$.

*Step 3: off-diagonal blocks.* Let $X \in \mathcal{S}_1$, $X \geq 0$, and $Z = PXP^\perp + P^\perp XP$. For real $t$, $(P + tP^\perp)X(P + tP^\perp) = PXP + tZ + t^2P^\perp XP^\perp \geq 0$; applying $T_1$ and Step 2 gives $PXP + t\,T_1Z \geq 0$, and dividing by $|t|$ and letting $t \to \pm\infty$ (the positive cone of $\mathcal{S}_1$ is trace-norm closed) gives $T_1Z = 0$. Since every $X \in \mathcal{S}_1$ is a difference of two positive ones and $X \mapsto Z$ is linear, $T_1Z = 0$ for every $X$.

Hence $T_1X = PXP$. The instrument $(T_0, T_1)$ has measured effect $1 - P$, so $T_0X = (1 - P)X(1 - P)$. Conversely, these maps are positive, sum to a trace-preserving map, and fix $F_1$ and $F_0$; an effect $f$ vanishing on all states supported in $P$ and in $P^\perp$ has $PfP = P^\perp fP^\perp = 0$, and then $f \geq 0$ forces $f = 0$ (a positive operator with zero diagonal blocks is zero). So (Id) and (Sh) hold. □

**Example B.1** (Repeatable and ideal, but not sharp)**.** Let $P, R, Q$ be rank-one projections on $\mathbb{C}^3$ with sum 1, and $T_1\rho = P\rho P + \frac{1}{2}\operatorname{Tr}(R\rho)P$, $T_0\rho = Q\rho Q + \frac{1}{2}\operatorname{Tr}(R\rho)Q$. These are completely positive with $\operatorname{Tr}(T_1\rho + T_0\rho) = \operatorname{Tr}\rho$, and the measured effect is $P + \frac{1}{2}R$. The faces are the single states $P$ and $Q$, which are fixed, so (Id) holds, and the instrument is repeatable. Sharpness fails: the state $R$ is perfectly distinguished from both faces by the effect $R$. And $T_1$ is not a Lüders map.

## B.2 Lemma 4.3: the real cyclic representation

Since $\varphi$ is real-linear on self-adjoint operators, $\varphi_s$ is real-linear on $A_0$ with $\varphi_s(X^*) = \varphi_s(X)$ and $\varphi_s(X^*X) = \varphi(X^*X) \geq 0$; so $\langle X, Y\rangle_\varphi = \varphi_s(X^*Y)$ is symmetric and positive semidefinite. For $X, Y \in A_0$, $Y^*(\|X\|^2 - X^*X)Y \geq 0$, so $\langle XY, XY\rangle_\varphi \leq \|X\|^2\langle Y, Y\rangle_\varphi$ (operator norm on $\mathcal{H}$). By Cauchy–Schwarz, left multiplication by $X$ preserves the null space and descends to a bounded operator $\pi(X)$ on the completion, with $\|\pi(X)\| \leq \|X\|$ and $\pi(X^*) = \pi(X)^*$, since $\langle X^*Y, Y'\rangle_\varphi = \varphi_s(Y^*XY') = \langle Y, XY'\rangle_\varphi$.

Each $\pi(P_i)$ is a symmetric idempotent. If $\dim_{\mathbb{K}} \mathcal{H} < \infty$, then $A_0$ and hence $\mathcal{H}_\varphi$ are finite-dimensional. □

## B.3 Rank preservation (used in Appendices D and E)

In the Schrödinger picture, Lemma 4.5 reads as follows for an atom $q$ (primitive in $V$): after $(t, q)$ the state is a multiple of $q$, later updates keep it a multiple of an atom (Lemma B.2), and a rank-one $x \geq 0$ has $\langle x, x\rangle = (\operatorname{tr} x)^2$.

**Lemma B.2** (Rank preservation)**.** *In a Euclidean Jordan algebra $V$, if $x \geq 0$ is a nonnegative multiple of a primitive idempotent and $a \in V$, then $U_a x$ is again a nonnegative multiple of a primitive idempotent (or 0). In particular $\langle U_a x, U_a x\rangle = (\operatorname{tr} U_a x)^2$.* *Not formalized.*

*Proof.* Let $x = \mu q$ with $q$ primitive, and $y = U_a x$, which is $\geq 0$ because $U_a$ is positive. By the fundamental formula $U_y = \mu^2 U_a U_q U_a$, and $U_q$ has rank one as a linear map, with range $\mathbb{R}q$ by (3.1); so $U_y$ has rank at most one. Write $y = \sum_k \lambda_k c_k$ over a Jordan frame $\{c_k\}$ (Section 3.1). If two of the $\lambda_k$, say $\lambda_1, \lambda_2$, are nonzero, then $U_y c_1 = \lambda_1^2 c_1$ and $U_y c_2 = \lambda_2^2 c_2$ are linearly independent, a contradiction. So $y = \lambda_1 c_1$ with $\lambda_1 \geq 0$, and $\langle y, y\rangle = \lambda_1^2 \operatorname{tr} c_1 = (\operatorname{tr} y)^2$. □

## B.4 Lemma 4.6: the formal word calculus

By (4.1), for a letter $q \in \{p_j, e - p_j\}$, $E(q \cdot X) = QE(X)Q$ and $\theta(q \cdot X) = U_q \theta(X)$. Put $\kappa_j = 2\,p_j \cdot + 2\,(e - p_j) \cdot - \mathrm{id}$; by (4.4), $E(\kappa_j X) = \Sigma_j E(X) \Sigma_j$ and $\theta(\kappa_j X) = U_{r_j} \theta(X)$. For $g = (g_1, \ldots, g_k)$ let $\kappa_g = \kappa_{g_1} \circ \cdots \circ \kappa_{g_k}$, $\bar{g} = (g_k, \ldots, g_1)$, $\Sigma_g = \Sigma_{g_1} \cdots \Sigma_{g_k}$ and $\mathcal{U}_g = U_{r_{g_1}} \cdots U_{r_{g_k}}$. Then $E(\kappa_g X) = \Sigma_g E(X) \Sigma_g^*$ and $\theta(\kappa_g X) = \mathcal{U}_g \theta(X)$; since $U_{r_j}^2 = \mathrm{id}$ and $\Sigma_j^2 = 1$, $\mathcal{U}_{\bar{g}} = \mathcal{U}_g^{-1}$ and $\Sigma_{\bar{g}} = \Sigma_g^*$.

For an orbit label $\alpha = (g, i)$ define

$$\sigma_\alpha = \kappa_g \circ (p_i \cdot) \circ \kappa_{\bar{g}}, \qquad \tau_\alpha = \kappa_g \circ \kappa_i \circ \kappa_{\bar{g}}.$$

Since $P_\alpha = \Sigma_g P_i \Sigma_g^*$, $\Sigma_\alpha = \Sigma_g \Sigma_i \Sigma_g^*$, and $\mathcal{U}_g U_{p_i} \mathcal{U}_g^{-1} = U_{q_\alpha}$, $\mathcal{U}_g U_{r_i} \mathcal{U}_g^{-1} = U_{r_\alpha}$ (an automorphism $\alpha$ satisfies $\alpha U_x \alpha^{-1} = U_{\alpha x}$),

$$\begin{aligned} E(\sigma_\alpha X) &= P_\alpha E(X) P_\alpha, & \theta(\sigma_\alpha X) &= U_{q_\alpha} \theta(X), \\ E(\tau_\alpha X) &= \Sigma_\alpha E(X) \Sigma_\alpha, & \theta(\tau_\alpha X) &= U_{r_\alpha} \theta(X). \end{aligned}$$

With $[\varnothing]$ the empty word ($E = 1$, $\theta = e$), $X = \sigma_{\alpha_1} \cdots \sigma_{\alpha_s}[\varnothing]$ has $E(X) = P_{\alpha_1} \cdots P_{\alpha_s} \cdots P_{\alpha_1}$ and $\theta(X) = U_{q_{\alpha_1}} \cdots U_{q_{\alpha_s}} e = U_{q_{\alpha_1}} \cdots U_{q_{\alpha_{s-1}}} q_{\alpha_s}$; $\tau_\alpha$ gives the second assertion. The model-side identities hold in every associative algebra generated by projections $P_j$. □

*The lifts used for the purity and linear relations.* The lift for a purity relation is $\sigma_\alpha \sigma_\beta \sigma_\alpha[\varnothing] - 2\lambda_{\alpha\beta}\, \sigma_\alpha \sigma_\beta[\varnothing] + \lambda_{\alpha\beta}^2\, \sigma_\alpha[\varnothing]$, with model operator $T_{\alpha\beta}^2$. The lift of $Z^2$ for a linear relation is, by (4.6),

$$c_0^2[\varnothing] + 2c_0 \sum_r c_r \sigma_{\alpha_r}[\varnothing] + \sum_{r,s} c_r c_s \Big( \sigma_{\alpha_r} \sigma_{\alpha_s}[\varnothing] + \tfrac{1}{4} \big( \sigma_{\alpha_s}[\varnothing] - \tau_{\alpha_r} \sigma_{\alpha_s}[\varnothing] \big) \Big),$$

whose effect is $c_0^2 e + 2c_0 \sum_r c_r q_{\alpha_r} + \sum_{r,s} c_r c_s \big(U_{q_{\alpha_r}} q_{\alpha_s} + \frac{1}{4}(q_{\alpha_s} - U_{r_{\alpha_r}} q_{\alpha_s})\big) = (c_0 e + \sum_r c_r q_{\alpha_r})^2$ by (4.5). Both $T_{\alpha\beta}$ and $Z$ are self-adjoint elements of $A_0$, as Lemma 4.4 requires.

### B.5 Lemma 4.7: the orbit spans $B$

Call the right side $B'$; it lies in $B$. Let $x \in B'$ and $\alpha = (g, i)$. Then $U_{q_\alpha} x = \lambda_{q_\alpha}(x) q_\alpha \in B'$ by (4.2). Also $U_{r_\alpha} = \mathcal{U}_g U_{r_i} \mathcal{U}_g^{-1}$ with $\mathcal{U}_g = U_{r_{g_1}} \cdots U_{r_{g_k}}$ (automorphisms conjugate quadratic maps), so $U_{r_\alpha}$ fixes $e$ and sends $q_\beta$ to the orbit point with label $((g, i, \bar{g}, h), k)$ for $\beta = (h, k)$, where $\bar{g} = (g_k, \ldots, g_1)$; so $U_{r_\alpha} x \in B'$. By (4.5), $q_\alpha \circ B' \subseteq B'$. Hence $B'$ is a unital subalgebra containing every $p_i$, and $B' = B$. □

### B.6 Proposition 4.8: details

*Intertwining.* For $\alpha = (g, i)$ and $\beta = (h, k)$ let $\beta' = ((g, i, \bar{g}, h), k)$. Then $q_{\beta'} = \mathcal{U}_g U_{r_i} \mathcal{U}_g^{-1} q_\beta = U_{r_\alpha} q_\beta$ and $P_{\beta'} = \Sigma_g \Sigma_i \Sigma_g^* \Sigma_h P_k \Sigma_h^* \Sigma_g \Sigma_i \Sigma_g^* = \Sigma_\alpha P_\beta \Sigma_\alpha$. With $U_{r_\alpha} e = e$ and well-definedness, $\eta(U_{r_\alpha} x) = \pi(\Sigma_\alpha)\eta(x)\pi(\Sigma_\alpha)$.

*Corner relation.* The purity relations give $\pi(P_\alpha)\pi(P_\beta)\pi(P_\alpha) = \lambda_{q_\alpha}(q_\beta)\pi(P_\alpha)$, and $\pi(P_\alpha)^2 = \pi(P_\alpha) = \lambda_{q_\alpha}(e)\pi(P_\alpha)$ since $\lambda_{q_\alpha}(e) = \operatorname{tr} q_\alpha / \langle q_\alpha, q_\alpha \rangle = 1$. By linearity, $\pi(P_\alpha)\eta(x)\pi(P_\alpha) = \lambda_{q_\alpha}(x)\pi(P_\alpha)$.

*Expectations.*

$$\langle \Omega, \pi(P_\alpha)\Omega \rangle = \varphi(P_\alpha) = \varphi(E(\sigma_\alpha[\varnothing])) = \langle \rho, \theta(\sigma_\alpha[\varnothing]) \rangle = \langle \rho, q_\alpha \rangle,$$

because $\varphi$ is linear and every word in the support of $\sigma_\alpha[\varnothing]$ is reproduced. □

### B.7 Theorem 4.11(b): general algebras

Let the tests be $B$-primitive and suppose a sequential projective representation reproduces all words while $\langle \rho, c_k \rangle > 0$ for some $k$.

*Normalising the tests.* A test with $p_i \in \{0, e\}$ is removed: the representation restricts to a representation of the remaining tests (words not containing test $i$), and $B$ does not change. If $p_i$ is not primitive in $B$, then $e - p_i$ is (Definition 4.10); exchange the outcomes of test $i$: replace $p_i$ by $e - p_i$ and $P_i$ by $1 - P_i$. Words correspond bijectively, with the same $E_w$ and $\theta(w)$, and $B$ does not change. Now every $p_i$ is primitive in $B$.

*Reconstruction.* Each $q_\alpha$ is primitive in $B$, so (4.2) holds, and Section 4.3 used nothing else. So Proposition 4.8 gives a unital Jordan homomorphism $\eta : B \to B(\mathcal{H}_\varphi)_{\mathrm{sa}}$ with $\langle \Omega, \eta(x)\Omega \rangle = \langle \rho, x \rangle$.

*A summand of positive weight.* $\langle \Omega, \eta(c_k)\Omega \rangle = \langle \rho, c_k \rangle > 0$, so $\eta(c_k) \neq 0$. Composed with an isomorphism $J \cong A_k$, the restriction of $\eta$ to $A_k$ is a Jordan homomorphism from $J$ into the associative algebra of bounded operators on $\mathcal{H}_\varphi$, so Theorem F.5 gives $\eta(c_k) = 0$, a contradiction. □

## B.8 The state extension in Section 4.4

*The pieces.* Since $c_{\mathrm{sp}} = e - \sum_k c_k$ is central in $B$, $x \mapsto U_{c_{\mathrm{sp}}} x = c_{\mathrm{sp}} \circ x$ is a surjective unital Jordan homomorphism $B \to B_{\mathrm{sp}}$; it preserves quadratic maps, which gives $\pi_{\mathrm{sp}}(\theta(w)) = \theta_{\mathrm{sp}}(w)$. For $x \in A_k$ with spectral decomposition $x = \sum \lambda c_\lambda$ in $A_k$ ($\sum c_\lambda = c_k$), $\|x\| c_k \pm x = \sum (\|x\| \pm \lambda) c_\lambda \geq 0$, which is the bound used in the discard step.

The faithful representations are $H_n(\mathbb{R}) \subset H_n(\mathbb{C})$; the symplectic $*$-embedding $M_n(\mathbb{H}) \to M_{2n}(\mathbb{C})$ restricted to $H_n(\mathbb{H})$; for a spin factor $\mathbb{R} \oplus \mathbb{R}^m$, $(t, v) \mapsto t\, 1 + \sum_k v_k \Psi_k$ with Hermitian matrices $\Psi_1, \ldots, \Psi_m$ satisfying $\Psi_k \Psi_l + \Psi_l \Psi_k = 2\delta_{kl}$ (tensor products of Pauli matrices); and $\mathbb{R} \subset \mathbb{C}$. A faithful unital Jordan homomorphism $\iota$ sends a spectral decomposition $x = \sum \lambda_k c_k$ to $\sum \lambda_k \iota(c_k)$ with nonzero orthogonal projections $\iota(c_k)$ summing to 1; so $\iota(x)$ has the spectrum of $x$, and $x \geq 0$ iff $\iota(x) \geq 0$.

*Extension.* On $H_d(\mathbb{C})$ the operator norm is the order-unit norm of 1. The functional $f$ on $\iota(B_{\mathrm{sp}})$ is positive with $f(1) = \langle \rho, c_{\mathrm{sp}} \rangle = 1$; from $-\|X\| 1 \leq X \leq \|X\| 1$, $|f(X)| \leq \|X\|$, so $\|f\| = 1$. A real Hahn–Banach extension $F$ has $\|F\| = 1$ and $F(1) = 1$. If $X \geq 0$ and $M = \|X\| > 0$, then $\|1 - X/M\| \leq 1$, so $F(1 - X/M) \leq 1$ and $F(X) \geq 0$. Hence $F$ is a state, $F(X) = \mathrm{Tr}(\sigma X)$ for a density matrix $\sigma$. □

## B.9 The generated algebras in Corollary 4.13

*Four atoms.* Let $A \cong J$ be a summand of $V$ with unit $c$, and $p_1, \ldots, p_4$ atoms of $A$ with $\mathrm{Jord}_A\langle c, p_1, \ldots, p_4 \rangle = A$ (they exist by Theorem 5.5). The subalgebra $A \oplus \mathbb{R}(e - c)$ contains $e$ and the $p_i$, so $B \subseteq A \oplus \mathbb{R}(e - c)$, and the image of $B$ under the Jordan homomorphism $x \mapsto c \circ x$ is a unital subalgebra of $A$ containing the $p_i$, hence $A$. If $e - c \notin B$ (and $c \neq e$), then $B \cap \mathbb{R}(e - c) = 0$, so $B$ is the graph $\{a + \chi(a)(e - c)\}$ of a linear $\chi : A \to \mathbb{R}$ which is multiplicative with $\chi(c) = 1$. Its kernel is an ideal of codimension one in the simple 27-dimensional algebra $A$, impossible. So $B = A \oplus \mathbb{R}(e - c)$ (or $B = A$ if $c = e$). The $p_i$ are primitive in $V$, hence in $B$, and $c/3$ is a state with $\langle c/3, c \rangle = 1$.

*Three tests.* Let $V = \bigoplus_k V_k$ be the decomposition into simple ideals and $\pi_k$ the projections. Then $x \mapsto (\pi_k(x))_k$ embeds $B$ in $\bigoplus_k \pi_k(B)$, and $\pi_k(B) = \mathrm{Jord}\langle e_k, \pi_k(p_1), \pi_k(p_2), \pi_k(p_3) \rangle$. If $V_k$ is special, so is $\pi_k(B)$. If $V_k \cong J$, each $\pi_k(p_i)$ is 0, $e_k$, an atom or the complement of an atom; after complementing, $\pi_k(B)$ is generated by at most three atoms, which lie in a complex subalgebra (Proposition 5.1), so $\pi_k(B)$ is special. A subalgebra of a direct sum of special algebras is special, and a special Euclidean Jordan algebra has no summand isomorphic to $J$ (that summand would be special). So Theorem 4.11(a) applies. Two tests are three tests with one repeated. □

## B.10 Proposition 4.14

*(a)* Given a representation over $\mathbb{K}$, the cyclic representation of Lemma 4.3 is a real one with a vector state and the same table, finite-dimensional if the original is (Appendix B.2). Conversely, by the same lemma it suffices to start from a real representation $(\mathcal{H}, P_i, \Omega)$ with a vector state, which is a state in the sense of

Definition 3.4. Extending scalars, $\mathcal{H} \otimes_{\mathbb{R}} \mathbb{K}$ with projections $P_i \otimes 1$ and the unit vector $\Omega \otimes 1$ is a representation over $\mathbb{K} = \mathbb{C}$ or $\mathbb{H}$ with the same table and the same finite dimension, since $\langle \Omega \otimes 1, (E_w \otimes 1)(\Omega \otimes 1) \rangle = \langle \Omega, E_w \Omega \rangle$.

*(b)* Spin factors are special, so Theorem 4.11(a) gives $\mathcal{J}(V) \subseteq \mathcal{Q}_{\text{fin}}(\mathbb{C})$. For strictness take three mutually orthogonal rank-one projections on $\mathbb{C}^3$ and the state $\mathrm{Tr}/3$: then $\Pr(\text{yes}_i) = \frac{1}{3}$ and $\Pr(\text{yes}_i, \text{yes}_j) = 0$ for $i \neq j$. In a spin factor (rank two, $\operatorname{tr} e = 2$) the first condition excludes $p_i \in \{0, e\}$, so each $p_i$ is primitive, and (3.1) gives $\Pr(\text{yes}_i, \text{yes}_j) = \langle \rho, p_i \rangle \langle p_i, p_j \rangle = \frac{1}{3} \langle p_i, p_j \rangle$. So $\langle p_i, p_j \rangle = 0$, and by Lemma 3.1 $p_j \in V_0(p_i) = \mathbb{R}(e - p_i)$, so $p_j = e - p_i$ for all $i \neq j$. Then $p_2 = e - p_1 = p_3$, contradicting $\langle p_2, p_3 \rangle = 0$.

*(c)* If $V$ is special, every $B$ is special and Theorem 4.11(a) gives finite-dimensional complex representations. If not, Corollary 4.13 gives four tests with no representation in any dimension. □

## B.11 Corollary 4.16: compactness

*Moment functionals.* Let $\mathcal{P}_n$ be the real unital $*$-algebra generated by symmetric idempotents $x_1, \ldots, x_n$ with no other relations; it is spanned by the words in the $x_i$. Call a linear $m : \mathcal{P}_n \to \mathbb{R}$ a *moment functional* if $m(1) = 1$, $m(Y^*) = m(Y)$ and $m(Y^*Y) \geq 0$ for all $Y$. A sequential projective representation gives one, $m(Y) = \varphi_s(Y(P_1, \ldots, P_n))$. Conversely, a moment functional gives a representation by the construction of Lemma 4.3: left multiplication by $x_i$ is bounded by 1 because $\langle Y, Y \rangle - \langle x_i Y, x_i Y \rangle = m\big(((1 - x_i)Y)^*(1 - x_i)Y\big) \geq 0$, so each $x_i$ acts as an orthogonal projection on the completion, and the vector state of $\Omega = [1]$ has $\|\pi(P_w)\Omega\|^2 = m(E_w)$.

By Cauchy–Schwarz $|m(u)| \leq 1$ for every word $u$. Since $x_i^2 = x_i$, the reduced words (no two equal adjacent letters) form a basis of $\mathcal{P}_n$, so a linear functional is the same as its vector of values on reduced words. So the set $\mathcal{M}_n$ of moment functionals, viewed in $\prod_u [-1, 1]$ over reduced words $u$, is convex and closed (each condition involves finitely many coordinates), hence compact by Tychonoff's theorem, and the word probabilities $m \mapsto m(E_w)$ are continuous.

*Finite length.* Let $\mathcal{R}$ be the compact convex set of admissible target states: all states of $J$ in (a); the states of $V$ with $\langle \rho, c_{\text{ex}} \rangle \geq \varepsilon$ in (b) (the set $\{\rho : \max_k \langle \rho, c_k \rangle \geq \varepsilon\}$, which is contained in it, is not convex when $m \geq 2$). For $L \geq 0$ let $M_L \subseteq \mathcal{M}_n \times \mathcal{R}$ be the set of pairs with $m(E_w) = \langle \rho, \theta(w) \rangle$ for all words of length at most $L$. Each $M_L$ is closed and $M_{L+1} \subseteq M_L$. If all were nonempty, their intersection would contain a pair agreeing on all words, an exact representation from an admissible state, which Theorem 4.1 (in (a)) or Theorem 4.11(b) (in (b): every admissible state gives positive weight to some $A_k$) excludes. So $M_L = \emptyset$ for some $L$.

Then the set $K_L$ of representable tables on words of length at most $L$ (the image of $\mathcal{M}_n$) and the set $T_L$ of target tables (the affine image of $\mathcal{R}$) are disjoint compact convex subsets of a finite-dimensional space, and strict separation gives an affine functional that is $\geq 0$ on $K_L$ and $\leq -\gamma < 0$ on $T_L$. Writing its constant term as a multiple of the empty-word probability makes it linear.

*No uniform constant $\gamma$ in (b) when $c_{\text{ex}} \neq e$.* Let $\rho_0 = (e - c_{\text{ex}})/\operatorname{tr}(e - c_{\text{ex}})$,

$\rho_1 = c_{\mathrm{ex}}/\operatorname{tr} c_{\mathrm{ex}}$ and $\rho_t = (1-t)\rho_0 + t\rho_1$. For $t > 0$, $\rho_t$ gives positive weight to an Albert summand; $\rho_0$ gives zero weight to all of them, so its table is representable by Theorem 4.11(a). A functional $\geq 0$ on representable tables and $\leq -\gamma$, for a fixed $\gamma > 0$, on the tables of all $\rho_t$, $t > 0$, would by continuity be $\leq -\gamma$ at $\rho_0$, a contradiction. □

# C Certificate details and conventions

## C.1 Lemma 6.4: reconstruction at finite length

*Proof of Lemma 6.4. Necessity.* Given a representation, take its real cyclic representation (Lemma 4.3) and $\mathcal{G}_{u,v} = \langle \pi(P_u)\Omega, \pi(P_v)\Omega \rangle$. It is positive semidefinite and (M1) is Lemma 4.3. (M2) holds because $P_{vi} + P_{v\bar{\imath}} = (P_i + 1 - P_i)P_v$. (M3) holds because $P_u^* P_v$ contains the factor $P_i(1-P_i) = 0$ where the last letters of $u$ and $v$ meet.

*Sufficiency.* Write $\mathcal{G}$ as the Gram matrix of vectors $e_u$, $u \in \mathcal{W}_L$, spanning a real Hilbert space $\mathcal{K}$ of dimension $\operatorname{rank}\mathcal{G}$. By (M2), $\|e_v - e_{vi} - e_{v\bar{\imath}}\|^2 = 0$, so $e_v = e_{vi} + e_{v\bar{\imath}}$ whenever $vi \in \mathcal{W}_L$. Let $E_i$ be the orthogonal projection onto $\mathcal{K}_i = \operatorname{span}\{e_u : u \text{ ends with } i\}$. The claim is that for $v \in \mathcal{W}_{L-1}$,

$$E_i e_v = \begin{cases} e_{vi} & \text{if the last letter of } v \text{ is not of test } i, \\ e_v & \text{if } v \text{ ends with } i, \\ 0 & \text{if } v \text{ ends with } \bar{\imath}. \end{cases} \tag{C.1}$$

In the first case $e_{vi} \in \mathcal{K}_i$ and $e_{v\bar{\imath}} \perp \mathcal{K}_i$ by (M3), so $E_i e_v = e_{vi}$. In the second $e_v \in \mathcal{K}_i$; in the third $e_v \perp \mathcal{K}_i$ by (M3). Now put $\Omega = e_\varnothing$, a unit vector by (M1). For a reduced word $w = (a_1, \dots, a_m)$, $m \leq L$, induction on (C.1) and $(1-E_i)e_v = e_v - E_i e_v$ give $Q_{a_m} \cdots Q_{a_1}\Omega = e_w$, with $Q_a = E_i$ for $a = i$ and $Q_a = 1 - E_i$ for $a = \bar{\imath}$. Hence the probability of $w$ in the representation $(\mathcal{K}, E_i, \langle \Omega, \cdot\, \Omega\rangle)$ is $\|e_w\|^2 = f(w)$. □

## C.2 The length-seven argument in detail

This subsection states and derives the Gram formulas of Lemma C.1 and gives the explicit vectors of Theorem 6.10. Hypothesis (H), walks, $L$ and the rule (6.2) are as in Section 6.2.

**The Gram formulas.** The vectors $a_x$, $\eta_s^H$, $\eta_s^C$, $\eta_s^R$, $\eta_{uv}^J$ are those of Section 6.2. For a fixed $s = (x, y, z)$ (the *row*) and an ordering $\rho$ of $B, C, D$ written as a word in the letters $x, y, z$, write $\mathrm{HH}(\rho) = \langle \eta_s^H, \eta_\rho^H \rangle$, and similarly CC, RR, $\mathrm{RH} = \langle \eta_s^R, \eta_\rho^H \rangle$ and $\mathrm{RC} = \langle \eta_s^R, \eta_\rho^C \rangle$; also $h_\rho = \mathrm{HH}(\rho)$. Each entry is read from (6.2): for instance $\mathrm{HH}(yxz) = L(xyzxy)$.

**Lemma C.1** (Gram formulas at length seven)**.** *Let $p_A, p_B, p_C, p_D$ be any four atoms of $J$ and assume (H). Let $s = (x, y, z)$ be any ordering of $B, C, D$; every formula below has the letters of $s$ substituted.*

*Narrower in Lean.* `LengthSeven.LemmaA.lemmaA` derives, for any four atoms, every Gram entry between the 28 path vectors, which include all $\eta_s^H$ and three each of the $\eta_s^C, \eta_s^R$. The printed formulas are not stated in this form.

*(a)* $L(xyz) = \omega_y$, *and* $\eta^C_{xyz} + \eta^C_{zyx} = 2\omega_y\,\xi, \quad \eta^R_{xyz} + \eta^R_{xzy} = 2R_{xyz}\,a_x$.

*(b)* $\langle \eta_s^H, \eta_\rho^C \rangle = \mathrm{CC}(\rho)$ *for every* $\rho$, *and*

$$\begin{aligned}
&\mathrm{HH}(xyz) = o_{Ax}o_{xy}o_{yz}, && \mathrm{HH}(xzy) = o_{Ax}o_{yz}R_{xyz},\\
&\mathrm{HH}(yxz) = 2R_{xyz}R_{Axy} - o_{xy}\omega_z, && \mathrm{HH}(yzx) = o_{xz}o_{yz}R_{Axy},\\
&\mathrm{HH}(zxy) = o_{xy}o_{yz}R_{Axz}, && \mathrm{HH}(zyx) = 2R_{xyz}\omega_y - o_{xy}o_{yz}R_{Axz};\\
&\mathrm{CC}(xyz) = o_{Ax}o_{xy}o_{yz}o_{Az}, && \mathrm{CC}(zyx) = 2\omega_y^2 - \mathrm{CC}(xyz),\\
&\mathrm{CC}(xzy) = 2o_{Ax}R_{Ayz}R_{xyz} - o_{Ax}o_{yz}\omega_x, && \mathrm{CC}(yzx) = 2\omega_y\omega_z - \mathrm{CC}(xzy),\\
&\mathrm{CC}(yxz) = o_{Az}h_{yxz}, && \mathrm{CC}(zxy) = 2\omega_x\omega_y - \mathrm{CC}(yxz);\\
&\mathrm{RR}(xyz) = o_{Ax}o_{xy}o_{yz}o_{xz}, && \mathrm{RR}(xzy) = 2o_{Ax}R^2_{xyz} - \mathrm{RR}(xyz),\\
&\mathrm{RR}(yzx) = o_{xy}h_{yzx}, && \mathrm{RR}(yxz) = 2R_{xyz}h_{yxz} - \mathrm{RR}(yzx),\\
&\mathrm{RR}(zxy) = o_{xz}h_{zxy}, && \mathrm{RR}(zyx) = 2o_{xz}R_{xyz}\omega_y - \mathrm{RR}(zxy);\\
&\mathrm{RH}(xyz) = \mathrm{RR}(xyz), && \mathrm{RH}(xzy) = \mathrm{RR}(xzy),\\
&\mathrm{RH}(yxz) = o_{xz}h_{yxz}, && \mathrm{RH}(yzx) = h_{yzx},\\
&\mathrm{RH}(zxy) = o_{xy}o_{xz}\omega_y, && \mathrm{RH}(zyx) = h_{zyx};\\
&\mathrm{RC}(xyz) = o_{Ax}o_{xy}o_{yz}R_{Axz}, && \mathrm{RC}(zyx) = 2o_{Ax}R_{xyz}\omega_y - \mathrm{RC}(xyz),\\
&\mathrm{RC}(yzx) = o_{Ax}h_{yzx}, && \mathrm{RC}(xzy) = 2o_{Ax}R_{xyz}\omega_z - \mathrm{RC}(yzx),\\
&\mathrm{RC}(zxy) = 2o_{Ax}R_{xyz}\omega_x - \mathrm{RC}(yxz),\\
&\mathrm{RC}(yxz) = 2R_{Axz}h_{yxz} - o_{xz}\big(2R_{Axy}\omega_y - o_{Ax}o_{xy}R_{Ayz}\big).
\end{aligned}$$

*(c) For distinct $u, v$ with third letter $w$, the inner products of $\eta^J_{uv}$ are, with $\eta^J_{ab}$: $uv$: $o_{Au}o_{uv}$; $vu$: $o_{uv}R_{Auv}$; $uw$: $o_{Au}R_{uvw}$; $wv$: $\omega_v$; $vw$: $o_{vw}R_{Auv}$; $wu$: $o_{uv}R_{Auw}$; with $\eta^H_\rho$: $uvw$: $o_{Au}o_{uv}o_{vw}$; $uwv$: $o_{Au}R_{uvw}$; $vuw$: $h_{yxz}$ at $(x, y, z) = (u, v, w)$; $vwu$: $o_{uv}\omega_w$; $wuv$: $o_{uv}R_{Auw}$; $wvu$: $o_{uv}\omega_v$; with $\eta^C_\rho$: $uvw$: $o_{Au}o_{uv}R_{Avw}$; $uwv$: $o_{Au}o_{Av}R_{uvw}$; $wuv$: $o_{Av}o_{uv}R_{Auw}$; $vuw$, $vwu$, $wvu$: $2\omega_u R_{Auv}$, $2\omega_w R_{Auv}$, $2\omega_v R_{Auv}$ minus the entries for $wuv$, $uwv$, $uvw$; with $\eta^R_\rho$: $uvw$: $o_{Au}o_{uv}R_{uvw}$; $uwv$: $2o_{Au}o_{uv}R_{uvw}$ minus the $uvw$ entry; $vuw$: $h_{yxz}$ at $(u, v, w)$; $vwu$: $2R_{uvw}R_{Auv}$ minus the $vuw$ entry; $wuv$: $o_{vw}o_{uv}R_{Auw}$; $wvu$: $2R_{uvw}\omega_v$ minus the $wuv$ entry.*

*Every word used has length at most six from $p_A$ (seven from $e/3$ after a leading "yes" on $A$); the words are those of Lemma 6.5 for the steps listed in Appendix C.2, together with all-"yes" words of length at most four and the words $(x, y, A)$, $(x, y, z, A)$, $(x, y, \bar{z}, A)$, $(x, y, z, x, A)$, $(x, y, z, \bar{x}, A)$ of two polarisations.*

**Reduction steps.** The two identities (6.4) hold under (H) by Lemma 6.5; the first, with history $t$ of length at most three, is written $[qv \mid t]$, the second, with $|t| \le 1$, is written $(q; vw \mid t)$. A step *from the left* is the same step applied to the reversed walk. The history lengths used below are at most two for $[\,\cdot\,]$ and at most one for $(\,\cdot\,)$.

**Four walk values.** For distinct $x, y, z \in \{B, C, D\}$:

(W1) $L(x) = o_{Ax}$, $L(xy) = R_{Axy}$, $L(xyx) = o_{Ax}o_{xy}$ and $L(xyzx) = o_{Ax}R_{xyz}$. The first and third are the probabilities of $(x)$ and $(x, y)$ from $p_A$. For the second, ($R2_0$) at $A$ on $\xi$ paired with $\xi$ gives $2L(xy) = L(xy) + L(yx) = 2R_{Axy}$. For the fourth, ($R2_0$) at $x$ on $a_x$ paired with $a_x$ gives $2L(xyzx) = 2R_{xyz}\|a_x\|^2$.

(W2) $L(xyz) = \omega_y$, and $\eta^C_{xyz} + \eta^C_{zyx} = 2\omega_y\xi$: this is ($R3_0$), since $P_AP_xP_yP_z\xi = \eta^C_{zyx}$; paired with $\xi$ it gives $2L(xyz) = 2\omega_y$.

(W3) $\eta^R_{xyz} + \eta^R_{xzy} = 2R_{xyz}a_x$: ($R2_0$) at $x$ on $a_x$.

(W4) $L(xyAzx) = o_{Ax}\omega_x$ and $L(xyAzyx) = L(xyzAyx) = o_{Ax}o_{xy}R_{Ayz}$. The first follows from ($R2_0$) at $A$ on $\xi$, the step $(x; Ay \,|\, \varnothing)$, $[Ax \,|\, \varnothing]$ and (W2), the terms in $R$ cancelling by (W1). The second is the polarisation $2\langle P_AY, P_zY\rangle = \|P_AY\|^2 + \|P_AP_zY\|^2 - \|P_A(1-P_z)Y\|^2$ with $Y = P_yP_x\xi$, which gives three word probabilities from $p_A$; the third walk then follows by $(y; zA \,|\, x)$ and (W1).

*Proof of Lemma C.1: method.* Part (a) is (W2) and (W3). The identity $\langle \eta^H_s, \eta^C_\rho\rangle = \mathrm{CC}(\rho)$ holds because $\eta^C_\rho = P_A\eta^C_\rho$ and $P_A$ is self-adjoint. Each other entry is the walk given by (6.2), reduced by one or two steps $[qv \,|\, t]$ or $(q; vw \,|\, t)$ to the values (W1)–(W4), to word probabilities, or to entries already obtained; the entries CC, RR, RC for the orderings $zyx$, $yzx$, $zxy$ and their analogues follow from (a) paired with the appropriate vector. Two examples. $\mathrm{HH}(yxz)$ is the walk $xyzxy$; the step $(x; yz \,|\, y)$ gives $2R_{xyz}L(xy) - L(xzyxy)$, and $[yx \,|\, \varnothing]$ gives $L(xzyxy) = o_{xy}L(xzy) = o_{xy}\omega_z$. $\mathrm{RC}(yxz)$ is the walk $xyzxAzxy$; the step $(x; Az \,|\, y)$ gives $2R_{Axz}h_{yxz} - L(xyzxzAxy)$, the step $[zx \,|\, xy]$ from the left gives $L(xyzxzAxy) = o_{xz}L(yxAzyx)$, and ($R2_0$) at $A$ replaces $P_AP_xP_y\xi$ by $2R_{Axy}\xi - P_AP_yP_x\xi$, so that $L(yxAzyx) = 2R_{Axy}\omega_y - L(xyAzyx)$, which is (W4). Every step uses ($R1_t$) with $|t| \le 2$, ($R2_t$) with $|t| \le 1$, ($R3_0$) only through (a), all-"yes" words of length at most four, or the two polarisations; so every word used has length at most six from $p_A$. □

*Proof of Theorem 6.10 for* M. Either hypothesis gives (H). Let $\mathrm{sgn}(s)$ be the sign of the permutation taking $(B, C, D)$ to $s$, put $\kappa_X = \sum_s \mathrm{sgn}(s)\,\eta^X_s$ for $X = H, C, R$, and

$$\Phi = 4\kappa_C + 10\kappa_R - 9\kappa_H.$$

For M, $o = \frac{1}{3}$, $R = \frac{1}{9}$ and $t_r = \omega_r = \frac{1}{27}$ for all labels. Lemma C.1 then gives the same values in every row; they are listed below, each line multiplied by the factor shown, for $\rho$ in the order $xyz, xzy, yxz, yzx, zxy, zyx$, whose signs relative to $s$ are $+, -, -, +, +, -$:

| | $xyz$ | $xzy$ | $yxz$ | $yzx$ | $zxy$ | $zyx$ |
|---|---|---|---|---|---|---|
| $\mathrm{HH} \times 243$ | 9 | 3 | 3 | 3 | 3 | −1 |
| $\mathrm{CC} \times 729$ | 9 | 3 | 3 | −1 | −1 | −7 |
| $\mathrm{RR} \times 729$ | 9 | −3 | −1 | 3 | 3 | −1 |
| $\mathrm{RH} \times 243$ | 3 | −1 | 1 | 3 | 1 | −1 |
| $\mathrm{RC} \times 729$ | 3 | −1 | 3 | 3 | −1 | −1 |

For example $\mathrm{HH}(yxz) = \frac{2}{81} - \frac{1}{81}$, $\mathrm{CC}(zyx) = \frac{2}{729} - \frac{9}{729}$ and $\mathrm{RC}(yxz) = \frac{2}{729} - \frac{1}{3}\left(\frac{2}{243} - \frac{3}{243}\right) = \frac{3}{729}$.

Because all traces of $\mathsf{M}$ are equal, the value of the entry for $(s, \rho)$ depends only on the relative ordering of $\rho$ with respect to $s$, and $\operatorname{sgn}(s)\operatorname{sgn}(\rho)$ is its sign. Hence $\langle \kappa_X, \kappa_Y \rangle = \sum_{s,\rho} \operatorname{sgn}(s)\operatorname{sgn}(\rho)\langle \eta_s^X, \eta_\rho^Y \rangle$ is six times the signed row sum. This is an identity about the actual vectors; no symmetry of the representation is assumed. With $\langle \kappa_H, \kappa_C \rangle = \|\kappa_C\|^2$ from (b) of the lemma,

$$\|\kappa_H\|^2 = \tfrac{60}{243}, \qquad \|\kappa_C\|^2 = \langle \kappa_H, \kappa_C \rangle = \tfrac{16}{243},$$
$$\|\kappa_R\|^2 = \tfrac{40}{243}, \qquad \langle \kappa_R, \kappa_H \rangle = \tfrac{48}{243}, \qquad \langle \kappa_R, \kappa_C \rangle = \tfrac{8}{243};$$

for instance $\|\kappa_H\|^2 = 6 \cdot \frac{9-3-3+3+3+1}{243}$. The Gram matrix of $(\kappa_C, \kappa_R, \kappa_H)$ is $\frac{1}{243}\begin{pmatrix} 16 & 8 & 16 \\ 8 & 40 & 48 \\ 16 & 48 & 60 \end{pmatrix}$. With the coefficients $4, 10, -9$, the terms $16 \cdot 16$, $100 \cdot 40$, $81 \cdot 60$, $2 \cdot 40 \cdot 8$, $-2 \cdot 36 \cdot 16$ and $-2 \cdot 90 \cdot 48$ give

$$\|\Phi\|^2 = \frac{256 + 4000 + 4860 + 640 - 1152 - 8640}{243} = -\frac{36}{243} = -\frac{4}{27} < 0.$$

$\Phi$ is a real combination of vectors of a real Hilbert space, so this is impossible. □

*Proof of Theorem 6.10 for* $\mathsf{S}$. For $\mathsf{S}$, $o_{AB} = \frac{1}{2}$, the other transition probabilities are $\frac{1}{3}$, $R_{ABC} = R_{ABD} = \frac{1}{6}$, $R_{ACD} = R_{BCD} = \frac{1}{9}$ and $\omega_1 = \omega_2 = \omega_3 = \frac{1}{18}$. Take

$$\Phi_{\mathsf{S}} = \sum_s \operatorname{sgn}(s)\left(\eta_s^C + w(s)\,\eta_s^R - 3\,\eta_s^H\right),$$

where $w(s) = 4$ if $B$ is the last letter of $s$ and $w(s) = 3$ otherwise. The traces of $\mathsf{S}$ are unchanged by exchanging the labels $C$ and $D$, while $\Phi_{\mathsf{S}}$ changes sign; so the rows $s$ and the image of $s$ under the exchange contribute equally to $\|\Phi_{\mathsf{S}}\|^2$, and only the rows $BCD$, $CBD$, $CDB$ need to be evaluated. Their values follow from Lemma C.1, and the next paragraph gives the contributions $\frac{4}{81}$, $\frac{1}{162}$, $-\frac{83}{1458}$ of these three rows, so $\|\Phi_{\mathsf{S}}\|^2 = 2\left(\frac{72+9-83}{1458}\right) = -\frac{2}{729} < 0$. □

**The configuration** $\mathsf{S}$. Write $(c_s^C, c_s^R, c_s^H)$ for the coefficients of $\eta_s^C, \eta_s^R, \eta_s^H$ in a vector $\sum_s (c_s^C \eta_s^C + c_s^R \eta_s^R + c_s^H \eta_s^H)$. Collecting each cross term in the row of its $\eta^C$ or $\eta^R$ vector, and using $\langle \eta_s^H, \eta_\rho^C \rangle = \mathrm{CC}(\rho)$, its squared norm is $\sum_s \Psi(s)$ with

$$\Psi(s) = \sum_\rho \left[ c_s^C c_\rho^C\,\mathrm{CC} + c_s^H c_\rho^H\,\mathrm{HH} + 2c_s^C c_\rho^H\,\mathrm{CC} + c_s^R c_\rho^R\,\mathrm{RR} + 2c_s^R c_\rho^H\,\mathrm{RH} + 2c_s^R c_\rho^C\,\mathrm{RC} \right],$$

the entries taken in row $s$ at $\rho$. For $\Phi_{\mathsf{S}}$, $(c^C, c^R, c^H)_s = \operatorname{sgn}(s)(1, w(s), -3)$, and the formulas of Lemma C.1 at the traces of $\mathsf{S}$ give $\Psi(BCD) = \frac{4}{81}$, $\Psi(CBD) = \frac{1}{162}$ and $\Psi(CDB) = -\frac{83}{1458}$; for example, in row $CDB$ the entry $\mathrm{HH}(yxz)$ is $2R_{CDB}R_{ACD} - o_{CD}\omega_B = \frac{2}{81} - \frac{1}{54} = \frac{1}{162}$. The images under the exchange of $C$ and $D$ have the same values. Hence $\|\Phi_{\mathsf{S}}\|^2 = 2\left(\frac{4}{81} + \frac{1}{162} - \frac{83}{1458}\right) = -\frac{2}{729}$, as used in the proof above.

## C.3 Proposition 6.17(a) and (b)

*Proof of Proposition 6.17(a), (b).* (a) If $E(c) = 0$, every representation gives $\varphi(\pi(E(c))) = 0$. Conversely, let $x = E(c) \neq 0$ and let $N$ bound the reduced lengths in its finite support. Let $B_N$ be the finite ball of reduced words of length at most $N$ in $\mathbb{Z}_2^{*n}$. For each generator $i$, swap $v$ with $r_i v$ whenever both belong to $B_N$, and fix $v$ when that edge leaves the ball. These are disjoint swaps and fixed points, hence an involutive permutation $s_i$ of $B_N$. Its real permutation matrix is symmetric, and the free-involution presentation gives a representation $\pi$. For a reduced word $g$ of length at most $N$, applying its letters from right to left to $\delta_e$ visits only reduced suffixes of $g$. Thus $\pi(g)\delta_e = \delta_g$, and $\pi(x)\delta_e$ is the nonzero coefficient vector of $x$. Since $x$ is fixed by reversal, $\pi(x)$ is a nonzero real symmetric matrix. A unit vector has nonzero quadratic expectation: a nonzero diagonal entry suffices, or, for a nonzero off-diagonal entry, one of the normalized vectors $\delta_a \pm \delta_b$ suffices. The projections $(I + s_i)/2$ and this vector state therefore give $\sum_w c_w \Pr(w) \neq 0$ in a finite real Hilbert space.

(b) $\bar{\omega}(f(c)) = E(c)$, so (a) applies; and $f(c) = 0$ in the free algebra gives $f(c)(p) = 0$ in every Jordan algebra. Arrow-of-time and repetition follow from $U_p e + U_{e-p} e = e$, $U_p^2 = U_p$ and $U_p U_{e-p} = 0$. For (6.11), (5.2) gives $U_{e-p_i} p_j - U_{p_i} p_j - p_j = -2p_i \circ p_j$, symmetric in $i, j$; and $U_{p_i} U_{e-p_j}(e-p_i) - U_{p_i} U_{p_j}(e-p_i) = U_{p_i}(e - p_i - 2p_j + 2p_i \circ p_j) = -2U_{p_i} p_j + 2U_{p_i}(p_i \circ p_j) = 0$, since $U_p L_p = U_p$. A classical process on $\{0,1\}^2$ started at $(1,1)$, in which test $i$ reads the first bit and resets the second to 0 and test $j$ reads the second bit, is repeatable and satisfies every arrow-of-time and repetition equality, but the first equality fails for it ($-1 \neq -2$). For the second, let test $i$ read the first bit and change nothing, and test $j$ read the second bit and set the first bit to 0 on outcome yes; this process is repeatable and gives $\Pr(i, j, \bar{\imath}) = 1 \neq 0 = \Pr(i, \bar{\jmath}, \bar{\imath})$ from $(1,1)$. □

## C.4 Identifying the configurations

Octonions are as in Appendix A.2 (Cayley–Dickson, $e_1 e_2 = e_3$, $e_1 e_4 = e_5$, $e_2 e_4 = e_6$, $e_3 e_4 = e_7$), and the atom of a vector $v$ with entries in an associative subalgebra is $p(v)_{ab} = v_a \bar{v}_b / |v|^2$ (Appendix A.4).

The vectors of $\mathsf{M}$, $\mathsf{S}$ and $\mathsf{Q}_\pm$ have entries in $\mathrm{span}\{1, i, j, \ell\}$, and the configuration is determined, up to an automorphism of $J$, by the requirement that $(i, j, \ell)$ be a basic triple. By fact (F4) of Appendix A.5, any two basic triples are related by an automorphism of $\mathbb{O}$, which acts entrywise as an automorphism of $J$ and changes no word probability. The same remark covers the other dyad convention $\bar{v}_a v_b$: it is the atom of the entrywise conjugate vector, and $(-i, -j, -\ell)$ is again a basic triple. The configuration $\mathsf{M}$ is also characterised without coordinates as the unique $F_4$-orbit of ordered quadruples on which $\mathcal{A}$ attains its maximum $\frac{4}{19683}$ (Proposition 5.12).

A configuration is confirmed by the invariants of Section 6: the thirteen traces, $\mathcal{A}$ and $\mathcal{D}$ printed there, and $\dim \mathrm{Jord}\langle e, p_A, \ldots, p_D \rangle = 27$. For $\mathsf{Q}_\pm$ the transition probabilities and shape invariants are those of the octonionic configuration and the double-pair traces are as stated after Proposition 6.11.

## C.5 The explicit $\mathbb{C}^{12}$ models

**The model of Proposition 6.16.** Let $\sigma_x, \sigma_y, \sigma_z$ be the Pauli matrices and

$$C_1 = \sqrt{-1}\,\sigma_x \otimes 1_2, \qquad C_2 = \sqrt{-1}\,\sigma_y \otimes 1_2, \qquad C_3 = \sqrt{-1}\,\sigma_z \otimes \sigma_x,$$

three anticommuting $4 \times 4$ matrices with square $-1$ and $C_1C_2C_3 = 1_2 \otimes \sigma_x$. On $\mathbb{C}^{12} = \mathbb{C}^4 \oplus \mathbb{C}^4 \oplus \mathbb{C}^4$ put

$$\mathcal{L}_C = \begin{pmatrix} 1_4 \\ -C_1 \\ 1_4 \end{pmatrix}, \qquad \mathcal{L}_D = \begin{pmatrix} 1_4 \\ -C_2 \\ -C_3 \end{pmatrix},$$

$$P_A = \begin{pmatrix} 1_4 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad P_B = \frac{1}{2}\begin{pmatrix} 1_4 & 1_4 & 0 \\ 1_4 & 1_4 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad P_C = \tfrac{1}{3}\mathcal{L}_C\mathcal{L}_C^\dagger, \quad P_D = \tfrac{1}{3}\mathcal{L}_D\mathcal{L}_D^\dagger,$$

with the state $\mathrm{Tr}/12$. All four are projections of rank four. The two eigenspaces of $1_2 \otimes \sigma_x$ in each $\mathbb{C}^4$ summand are the two sectors described after Proposition 6.16.

**The model of Proposition 6.11.** Represent a quaternion $q = q_0 + q_1 i + q_2 j + q_3\, ij$ by

$$\psi(q) = \begin{pmatrix} a_1 & -a_2 \\ \bar{a}_2 & \bar{a}_1 \end{pmatrix}, \qquad a_1 = q_0 + q_1\sqrt{-1}, \quad a_2 = q_2 + q_3\sqrt{-1},$$

a multiplicative map with $\psi(\bar{q}) = \psi(q)^\dagger$. For a quaternionic vector $v = (v_1, v_2, v_3)$ let $\Phi_v$ be the $6 \times 2$ matrix with blocks $\psi(v_1), \psi(v_2), \psi(v_3)$, divided by $|v|$; then $\Phi_v\Phi_v^\dagger = \iota_{\mathbb{H}}(p(v))$ is a projection of rank two. Explicitly $\psi(i) = \mathrm{diag}(\sqrt{-1}, -\sqrt{-1})$, $\psi(j) = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $\psi\big((i+j)/\sqrt{2}\big) = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} \sqrt{-1} & -1 \\ 1 & -\sqrt{-1} \end{smallmatrix}\right)$. For $\mathsf{M}$,

$$\Phi_A = \begin{pmatrix} 1_2 \\ 0 \\ 0 \end{pmatrix}, \quad \Phi_B = \frac{1}{\sqrt{3}}\begin{pmatrix} 1_2 \\ \sqrt{2}\,1_2 \\ 0 \end{pmatrix}, \quad \Phi_C = \frac{1}{\sqrt{3}}\begin{pmatrix} 1_2 \\ \psi(i) \\ 1_2 \end{pmatrix}, \quad \Phi_D^\pm = \frac{1}{\sqrt{3}}\begin{pmatrix} 1_2 \\ \psi(j) \\ \pm\psi\left(\frac{i+j}{\sqrt{2}}\right) \end{pmatrix};$$

for $\mathsf{S}$ replace $\Phi_B$ by $\frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1_2 \\ 1_2 \\ 0 \end{smallmatrix}\right)$. The model on $\mathbb{C}^{12} = \mathbb{C}^6 \oplus \mathbb{C}^6$ is

$$P_X = \Phi_X\Phi_X^\dagger \oplus \Phi_X\Phi_X^\dagger \quad (X = A, B, C), \qquad P_D = \Phi_D^+\Phi_D^{+\dagger} \oplus \Phi_D^-\Phi_D^{-\dagger}, \qquad \varphi = \mathrm{Tr}/12,$$

four projections of rank four with entries in $\mathbb{Q}(\sqrt{2}, \sqrt{-1})$. The first summand is the standard representation of $\mathsf{Q}_+$, the second that of $\mathsf{Q}_-$.

**A rational form for the M model.** The atoms (5.13) of $\mathsf{M}$ have entries in $\mathbb{Q}(\sqrt{2})$ rather than $\mathbb{Q}$, but they lie in another rational form of the Albert algebra. With $k = ij$, let $\gamma$ act on $\mathbb{H}$ by conjugation by $k$, extended coefficientwise to $\mathbb{H} \oplus \mathbb{H}\ell$; it is an automorphism negating $i, j$ and fixing $k, \ell$. For $D = \mathrm{diag}(1, -1, 1)$, $\beta(X) = D\gamma(X)D$ is an involutive rational Jordan automorphism with eigenspaces $J_+, J_-$ of dimensions

15 and 12 in the rational-coordinate Albert algebra, and $J_{\mathrm{rat}} = J_+ \oplus \sqrt{2}J_-$ is a 27-dimensional rational form, closed under the Jordan product. With $F_{ab}(z)$ denoting $z$ in the upper slot and its conjugate in the lower slot, it contains the atoms

$$p_A = E_{11}, \qquad p_B = (E_{11} + 2E_{22} + \sqrt{2}F_{12}(1))/3,$$
$$p_C = (e - F_{12}(i) + F_{13}(1) + F_{23}(i))/3, \qquad p_D = (e - F_{12}(j) - F_{13}(\ell) - F_{23}(j\ell))/3.$$

In a basis adapted to $J_{\mathrm{rat}}$ all multiplication and quadratic maps have rational matrices, so all branch probabilities from $e/3$ or $p_A$ are rational.

## C.6 Six-step models for all generating quadruples

The construction first gives the two quaternionic blocks; their mixture is then shown to match every word through length six. Throughout, the initial state is $e/3$ and $\mathcal{A} > 0$. The notation $\tau(o, R)$ denotes the right side of the trace relation (5.3).

**Proposition C.2** (Quaternionic pairs on a line of double-pair traces)**.** *Let $p_A, p_B, p_C, p_D$ be atoms of $J$ with $\mathcal{A} > 0$, with transition probabilities $o$, shape invariants $R$ and double-pair traces $t^J = (t_1, t_2, t_3)$. Let $\Pi$ be the plane $\{t \in \mathbb{R}^3 : t_1+t_2+t_3 = \tau(o, R)\}$, and for $t \in \Pi$ let $M(t)$ and $\mathcal{D}(t)$ be the matrix and statistic of (5.9), computed at the base $A$ from $(o, R, t)$.*

*(a) $t^J \in \Pi$, and $M(t^J)$ is positive definite.*

*(b) If $t \in \Pi$, $M(t) \geq 0$ and $\mathcal{D}(t) = 0$, then there are four atoms of a unital subalgebra $H_3(\mathbb{H}) \subset J$ whose thirteen traces are $(o, R, t)$.*

*(c) For every line $\Pi_1 \subset \Pi$ through $t^J$ there are $t^{\pm} \in \Pi_1$ satisfying the hypotheses of (b), with $t^J = \lambda t^+ + (1 - \lambda)t^-$ for some $\lambda \in (0, 1)$. The $\lambda$-mixture of the standard representations on $\mathbb{C}^6$ of the two quaternionic quadruples of (b) is a sequential projective representation on $\mathbb{C}^{12}$ whose thirteen traces from $e/3$ are those of $p_A, \ldots, p_D$, and whose two blocks each have the octonionic $o$ and $R$.*

*Formalized.* `MatchingModels.prop_cer_quatpair_a`, `MatchingModels.prop_cer_quatpair_b`, `MatchingModels.prop_cer_quatpair_c_line`; the mixture on $\mathbb{C}^{12}$: `MatchingModels.prop_cer_universalsix_witness`.

*Proof.* (a) The first statement is Proposition 5.3. For the second, $\mathcal{A} > 0$ makes all transition probabilities positive, and the proof of Proposition 5.10 gives the signed diagonal congruence $M = TG'T$ with $T$ invertible, where $G'$ is, by the proof of Theorem 5.5 (step "Traces become octonion data"), the Gram matrix of three imaginary octonions. So $M \geq 0$. Moreover $\det M = \mathcal{D} + (c_4 - S)^2 \geq \mathcal{D} = \mathcal{A}/\prod o_{ij} > 0$, so $M > 0$.

(c) The two endpoints come from a chord of complex density matrices; part (b) is not needed. By Proposition 5.6, $\Delta_W(p_A, p_B, p_C) \neq 0$, so $p_A, p_B, p_C$ generate a complex subalgebra $W$; let $\psi : W \to H_3(\mathbb{C})$ be a unital Jordan isomorphism, which preserves the trace form. Put $\sigma = \psi(P_W p_D)$, a positive definite density matrix by Corollary 5.7(iii). Every trace involving $p_D$ involves it once, and equals $\langle p_D, x\rangle = \mathrm{Tr}(\sigma\, \psi(x))$ for $x$ one of the nine elements of $W$ listed in the paragraph after this proof, with $(X, Y, Z) = (p_A, p_B, p_C)$: the transition probabilities and shape invariants pair $\sigma$ with the first six, the double-pair traces with the last three. Those nine elements form a basis of $W$. So the traceless Hermitian matrices $H$

trace-orthogonal to the images of the first six form a space $\mathcal{H}$ of dimension at least two, and $H \mapsto (\mathrm{Tr}(H\psi(x)))_x$, over the last three elements $x$, is injective on $\mathcal{H}$.

For $0 \neq H \in \mathcal{H}$, the $s$ with $\sigma + sH \geq 0$ form a closed bounded interval with 0 in its interior ($\sigma > 0$, and $H$ has eigenvalues of both signs). At its ends $\sigma_\pm$ has rank at most two and trace one, so $\sigma_\pm = xx^* + yy^*$ with $x, y \in \mathbb{C}^3$ ($y = 0$ allowed). The unit quaternionic column $v = x + yj$ has $vv^*$ with complex part $\sigma_\pm$ (end of Section 5.3). The atoms $\psi(p_A)$, $\psi(p_B)$, $\psi(p_C)$, $vv^*$ of $H_3(\mathbb{H}) \supset H_3(\mathbb{C})$, placed in $J$ by a unital embedding, therefore have the transition probabilities and shape invariants of $p_A, \ldots, p_D$, and double-pair traces $t^\pm$ with $t^J = \lambda t^+ + (1-\lambda)t^-$, where $\sigma = \lambda\sigma_+ + (1-\lambda)\sigma_-$ and $\lambda \in (0,1)$. These atoms lie in a quaternionic subalgebra, so $\mathcal{A} = 0$ and $\mathcal{D}(t^\pm) = 0$ (Theorem 5.5, Proposition 5.10); they satisfy the trace relation, so $t^\pm \in \Pi$; and the argument of (a) gives $M(t^\pm) \geq 0$. As $H$ ranges over $\mathcal{H}$, the directions $t^\pm - t^J$ are the images of $H$ under an injective linear map into the two-dimensional direction space of $\Pi$; so $\dim \mathcal{H} = 2$ and every line $\Pi_1 \subset \Pi$ through $t^J$ arises.

The standard representation of $H_3(\mathbb{H})$ on $\mathbb{C}^6$ reproduces the table of each quaternionic quadruple from $e/3$ (with the state Tr/6), so the direct sum on $\mathbb{C}^{12}$ with state $\lambda\,\mathrm{Tr}/6 \oplus (1-\lambda)\,\mathrm{Tr}/6$ has table $\lambda f_+ + (1-\lambda) f_-$. Transition probabilities, shape invariants and double-pair traces are affine functions of the table, so the mixture has $o$, $R$ and $\lambda t^+ + (1-\lambda)t^- = t^J$.

(b) This converse is not used for Proposition 6.2. Write $a = \sqrt{o_{AB}}$, $b = \sqrt{o_{AC}}$, $c = \sqrt{o_{AD}}$, $d = \sqrt{o_{BC}}$, $e' = \sqrt{o_{BD}}$, $f = \sqrt{o_{CD}}$, and let $c_1, c_2, c_3, c_4, S$ be as in (5.9). Since $M \geq 0$ is a real symmetric $3 \times 3$ matrix, there are $u_1, u_2, u_3 \in \mathbb{R}^3 = \mathrm{Im}\,\mathbb{H}$ with Gram matrix $M$. Replacing all three by their negatives if necessary, it may be assumed that $\det(u_1, u_2, u_3) = S - c_4$; this is possible because $(c_4 - S)^2 = \det M = \det(u_1, u_2, u_3)^2$ by $\mathcal{D} = 0$. Put $h_a = c_a + u_a$, unit quaternions since $|u_a|^2 = M_{aa} = 1 - c_a^2$, and let $\Theta$ be the quaternionic Hermitian matrix, rows and columns indexed by $A, B, C, D$,

$$\Theta = \begin{pmatrix} 1 & a & b & c \\ a & 1 & d\,h_1 & e'\bar{h}_3 \\ b & d\,\bar{h}_1 & 1 & f\,h_2 \\ c & e'h_3 & f\,\bar{h}_2 & 1 \end{pmatrix}.$$

For unit vectors $v_i \in \mathbb{H}^n$ with Gram matrix $\Theta_{ij} = v_i^* v_j$, the atoms $v_i v_i^*$ of $H_n(\mathbb{H})$ have $o_{ij} = |\Theta_{ij}|^2$, $R_{ijk} = \mathrm{Re}(\Theta_{ij}\Theta_{jk}\Theta_{ki})$ and $Q_{i;jkl} = \mathrm{Re}(\Theta_{ij}\Theta_{jk}\Theta_{kl}\Theta_{li})$, since the real trace is cyclic. For this $\Theta$ this gives the prescribed $o_{ij}$; $R_{ABC} = abd\,c_1$, $R_{ACD} = bcf\,c_2$, $R_{ABD} = ace'c_3$ and $R_{BCD} = de'f\,\mathrm{Re}(h_1h_2h_3)$; and $Q_{A;BCD} = adfc\,\mathrm{Re}(h_1h_2)$ and its two cyclic companions. Now $\mathrm{Re}(h_ah_b) = c_ac_b - \langle u_a, u_b\rangle = c_ac_b - M_{ab}$, which is the number $K_{ab}$ of Section 5.4, and $\mathrm{Re}(h_1h_2h_3) = S + \mathrm{Re}(u_1u_2u_3) = S - \det(u_1, u_2, u_3) = c_4$. Since $(Q_{A;BCD}, Q_{A;CDB}, Q_{A;DBC}) = (t_1 + t_3 - t_2,\ t_1 + t_2 - t_3,\ t_2 + t_3 - t_1)$ is an invertible linear function of $t$, all thirteen traces computed from $\Theta$ are $(o, R, t)$.

It remains to see that $\Theta \geq 0$ with quaternionic rank three, so that $\Theta$ is the Gram matrix of four unit vectors in $\mathbb{H}^3$. Its leading $3 \times 3$ block $\Theta_3$ has entries in the commutative field $\mathbb{R} + \mathbb{R}u_1 \cong \mathbb{C}$, and its leading principal minors are 1, $1 - o_{AB} > 0$ and $N(p_A + p_B + p_C) = 1 - o_{AB} - o_{AC} - o_{BC} + 2R_{ABC}$, which is positive because $\mathcal{A} > 0$ forces $\Delta_W(p_A, p_B, p_C) \neq 0$ (Proposition 5.6). So $\Theta_3 > 0$, and with

$g = (c,\ e'\bar{h}_3,\ f h_2)^{\mathsf{T}}$ the Schur complement $s = 1 - g^* \Theta_3^{-1} g$ is real and $\Theta$ is congruent to $\operatorname{diag}(\Theta_3, s)$.

The claim is that for unit vectors $v_A, \dots, v_D \in \mathbb{H}^4$ with a positive definite Gram matrix $\Theta$ of this shape,

$$t_1 + t_2 + t_3 - \tau(o, R) = -\tfrac{1}{2} N(p_A + p_B + p_C)\, s, \tag{C.2}$$

with the traces computed from $\Theta$ as above. This is the quaternionic form of Remark 5.4(b).

Let $\iota_{\mathbb{H}} : M_4(\mathbb{H}) \to M_8(\mathbb{C})$ be the standard embedding. A matrix $x \in H_4(\mathbb{H})$ has four real eigenvalues $\mu_1, \dots, \mu_4$ with $\operatorname{Re}\operatorname{Tr}(x^k) = \sum_i \mu_i^k$, and $\iota_{\mathbb{H}}(x)$ has the same eigenvalues, each twice. So $e_4(x) = \mu_1\mu_2\mu_3\mu_4$ is a polynomial in the entries of $x$, by Newton's identities, with $e_4(x)^2 = \det \iota_{\mathbb{H}}(x)$ and $e_4(x) > 0$ for $x > 0$.

Run the proof of Proposition 5.3 for $x = \sum_i \lambda_i v_i v_i^*$. Newton's identity for four eigenvalues is the one used there with the extra term $-4e_4$, so comparing the coefficients of $\lambda_A\lambda_B\lambda_C\lambda_D$ gives $8\,(t_1 + t_2 + t_3 - \tau(o, R)) = -4m_4$, where $m_4$ is the coefficient of $\lambda_A\lambda_B\lambda_C\lambda_D$ in $e_4(x)$. With $V = (v_A, \dots, v_D)$ and $\Lambda = \operatorname{diag}(\lambda_i)$ the identities are $x = V\Lambda V^*$ and $\det \iota_{\mathbb{H}}(x) = (\det \Lambda)^2 \det \iota_{\mathbb{H}}(\Theta)$. Hence $e_4(x) = \lambda_A\lambda_B\lambda_C\lambda_D \det \iota_{\mathbb{H}}(\Theta)^{1/2}$ for positive $\lambda_i$, so identically, and $m_4 = \det \iota_{\mathbb{H}}(\Theta)^{1/2}$.

The congruence of $\Theta$ with $\operatorname{diag}(\Theta_3, s)$ is by a unitriangular matrix, whose image under $\iota_{\mathbb{H}}$ has determinant one, so $\det \iota_{\mathbb{H}}(\Theta) = \det \iota_{\mathbb{H}}(\Theta_3)\, s^2$. Here $\det \iota_{\mathbb{H}}(\Theta_3) = (\det \Theta_3)^2$ because $\Theta_3$ has entries in $\mathbb{R} + \mathbb{R}u_1$, and $\det \Theta_3 = N(p_A + p_B + p_C)$. Since $\Theta > 0$ gives $s > 0$, this is (C.2).

Both sides of (C.2) are polynomials in the entries of $\Theta$ (on the right, $\det \Theta_3\, s = \det \Theta_3 - g^* \operatorname{adj}(\Theta_3) g$) on the real affine space of quaternionic Hermitian matrices with unit diagonal whose leading block has entries in $\mathbb{R} + \mathbb{R}u_1$. They agree on the nonempty open subset where $\Theta > 0$, hence on the whole space.

For this $\Theta$ the traces lie in $\Pi$, so the left side is zero and $s = 0$. Hence $\Theta \geq 0$ has rank three and is the Gram matrix of unit vectors $v_A, \dots, v_D \in \mathbb{H}^3$; the atoms $v_i v_i^*$ of $H_3(\mathbb{H})$, placed in $J$ by a unital embedding, have the traces $(o, R, t)$. In words: $s = 0$ says that the vector with the prescribed inner products with $v_A, v_B, v_C$ has norm one, and the plane $\Pi$ is exactly this unit-norm condition. □

**Traces with one occurrence of an atom.** Let $X, Y, Z, D$ be the four atoms, and let a Jordan monomial contain $D$ exactly once. Associativity of the trace form moves that occurrence to the outside, so its trace is $\langle D, \Phi \rangle$ for a Jordan polynomial $\Phi$ in $e, X, Y, Z$. Since $\mathcal{A} > 0$, Proposition 5.6 gives $\Delta_W(X, Y, Z) \neq 0$. Thus $X, Y, Z$ generate a copy of $H_3(\mathbb{C})$, and the following nine elements form a real basis of it:

$$X, Y, Z, \quad X \circ Y, X \circ Z, Y \circ Z, \quad (X \circ Y) \circ Z, (Y \circ Z) \circ X, (Z \circ X) \circ Y.$$

These are also the nine monomials in the universal algebra of three primitive axes of Jordan type $\frac{1}{2}$. Gorshkov and Staroletov give its multiplication table and the determinant of the Frobenius form on them[181]. Indeed, write $X_i = a_i a_i^*$ with unit columns $a_i \in \mathbb{C}^3$, $g_{ij} = a_i^* a_j$ and $b = g_{12}g_{23}g_{31}$; then $\Delta_W = |\det(a_1, a_2, a_3)|^2 (\operatorname{Im} b)^2 \neq$

[181] Gorshkov and Staroletov, *Primitive 3-generated axial algebras*, note 139. Props. 5, 7.

0 makes the columns independent and all $g_{ij}$ nonzero, and a direct calculation in the coordinates $a_i a_j^*$ shows that the nine elements are independent (the last three give, modulo the first six and after rescaling, the coordinate rows $(0,1,-1)$, $(-1,1,0)$, $(-1,0,-1)$, of determinant $-2$).

The coefficients of $\Phi$ in this basis depend only on the three transition probabilities and the shape invariant of $X, Y, Z$. Indeed, these data determine their complex Gram matrix up to phases and complex conjugation: they determine $|g_{ij}|$ and $\operatorname{Re} b$, while $|b|$ is their fixed product. Thus they determine the triple up to a Jordan isomorphism, which carries both $\Phi$ and the displayed basis to their counterparts. Pairing the first six basis elements with $D$ gives transition probabilities and shape invariants; pairing the last three gives the three double-pair traces. Consequently the trace of the monomial is the same function of $(o, R, t)$ for the Albert and quaternionic configurations, affine in $t$ when $(o, R)$ is fixed. A factor involving at most three distinct atoms depends only on their $(o, R)$, by the same complex-triple argument. Every triple used here is nondegenerate, because the blocks preserve $(o, R)$ and hence the nonzero $\Delta_W$ of each Albert triple.

**Expansion and counting.** For a chronological word $w = (q_1, \dots, q_m)$, expand $\Pr_{e/3}(w) = \frac{1}{3}\operatorname{tr}(U_{q_m} \cdots U_{q_1} e)$ using

$$U_p x = \langle p, x \rangle\, p, \qquad U_{e-p} = \mathrm{id} - 2L_p + U_p.$$

Each chosen $U_p$ splits a term into scalar trace factors. If the split positions are $s_1 < \cdots < s_r$, these factors use the intervals $[1, s_1], [s_1, s_2], \dots, [s_r, m]$; neighbouring intervals share only their split position. The internal factors have the form $\langle p_{s_j}, L_{c_1} \cdots L_{c_k} p_{s_{j+1}} \rangle$, with the corresponding end factors involving $e$ and the trace. If there is no split there is just one trace factor. Identity terms omit positions. Thus each factor is the trace of a Jordan monomial using each position of its interval at most once, and the formal expansion and its rational coefficients are the same in $J$ and $H_3(\mathbb{H})$.

A factor with at most seven positions containing all four atoms has some atom only once: two occurrences of each would require eight positions. The preceding argument makes it affine in $t$. Two factors each containing all four atoms require at least $4 + 4 - 1 = 7$ positions, because their intervals share at most one position. Hence for $m \le 6$ every term has at most one factor depending on $t$, affinely, and all its other factors depend only on $(o, R)$. This proves that every such word probability is one universal affine function of $t$ at fixed $(o, R)$.

**The representation.** Apply Proposition C.2(c), so $t^J = \lambda t^+ + (1-\lambda) t^-$, and write $p_i^\pm$ for the atoms of the two quaternionic quadruples, with tables $\Pr^\pm$. Affinity gives, for every $|w| \le 6$,

$$\underset{e/3}{\overset{J}{\Pr}}(w) = \lambda \underset{e/3}{\overset{+}{\Pr}}(w) + (1-\lambda) \underset{e/3}{\overset{-}{\Pr}}(w).$$

Under the standard embedding $\iota_{\mathbb{H}} : H_3(\mathbb{H}) \to H_6(\mathbb{C})$, the projections $P_i = \iota_{\mathbb{H}}(p_i^+) \oplus \iota_{\mathbb{H}}(p_i^-)$ and density matrix $\rho = \lambda I_6/6 \oplus (1-\lambda) I_6/6$ have precisely this table, since $\mathrm{Tr} \circ \iota_{\mathbb{H}} = 2\,\mathrm{tr}$. This proves Proposition 6.2.

# D The spinor realization

This appendix proves Theorem 8.1. Let $p_1, \dots, p_n$ be atoms of $J$ and $\rho$ a state. Write $u_i = e - p_i$ for the unit of $J_0(p_i)$, and $T_{ji}$ for the restriction of $U_{e-p_j}$ to $J_0(p_i)$, a linear map $J_0(p_i) \to J_0(p_j)$.

**Peirce facts.** For an atom $p$ the Peirce-$\frac{1}{2}$ component of any $x$ is traceless, so, by (3.1),

$$\operatorname{tr} x = \operatorname{tr} U_p x + \operatorname{tr} U_{e-p} x = \langle p, x \rangle + \operatorname{tr} U_{e-p} x. \tag{D.1}$$

**Coordinates on a Peirce space.** $J_0(p) \cong H_2(\mathbb{O})$ has dimension ten. In a frame with $p = E_{11}$, a traceless element of $J_0(p)$ is $v = \operatorname{diag}(0, a, -a) + F_{23}(z)$, and $v \circ v = (a^2 + |z|^2)u$, $\langle v, v \rangle = 2(a^2 + |z|^2)$; so $v \circ v = \frac{1}{2}\langle v, v \rangle\, u$ with $u = e - p$. Choose an orthonormal basis $f_1, \dots, f_9$ of the traceless part and put $e_k = \sqrt{2} f_k$. Then $e_k \circ e_l = \delta_{kl} u$, and

$$\varphi_p : \mathbb{R}^{1,9} \to J_0(p), \qquad (x_0, \mathbf{x}) \mapsto x_0 u + \textstyle\sum_k x_k e_k,$$

is a Jordan isomorphism from the spin factor $\mathbb{R} \oplus \mathbb{R}^9$ (Section 3.1). Write $q_M(x) = x_0^2 - |\mathbf{x}|^2$. The element $y = \varphi_p(x)$ has eigenvalues $x_0 \pm |\mathbf{x}|$ on two orthogonal atoms of $J_0(p)$ and $0$ on $p$; it is positive exactly when $x$ lies in the forward cone $x_0 \geq |\mathbf{x}|$, and $\operatorname{tr} y = 2x_0$.

**The Freudenthal adjoint.** For $x = \sum_k \lambda_k c_k$ in a Jordan frame, the adjoint $x^\# = x^{\circ 2} - \operatorname{tr}(x)\, x + e_2(x)\, e$, with $e_2$ the second elementary symmetric function of the eigenvalues, is $\lambda_2\lambda_3 c_1 + \lambda_1\lambda_3 c_2 + \lambda_1\lambda_2 c_3$. Hence, for $y = \varphi_p(x)$,

$$y^\# = q_M(x)\, p, \tag{D.2}$$

and $(e - p)^\# = p$. The adjoint satisfies the composition law

$$(U_a x)^\# = U_{a^\#}\, x^\# \qquad (a, x \in J), \tag{D.3}$$

a polynomial identity valid in every cubic Jordan algebra, invertible or not (McCrimmon[182]).

**The similitude.** Let $y \in J_0(p_i)$ and write $q_i(y) = q_M(\varphi_{p_i}^{-1} y)$. By (D.3) with $a = e - p_j$, then (D.2),

$$(T_{ji} y)^\# = U_{p_j}\big(q_i(y)\, p_i\big) = o_{ij}\, q_i(y)\, p_j.$$

Since $T_{ji} y \in J_0(p_j)$, (D.2) also gives $(T_{ji} y)^\# = q_j(T_{ji} y)\, p_j$. Hence

$$q_j(T_{ji} y) = o_{ij}\, q_i(y). \tag{D.4}$$

In coordinates, $M_{ji} = \varphi_{p_j}^{-1} T_{ji} \varphi_{p_i}$ satisfies $M_{ji}^{\mathsf{T}} \eta M_{ji} = o_{ij} \eta$, where $\eta$ is the diagonal matrix with entries $1, -1, \dots, -1$. If $o_{ij} > 0$, $M_{ji}$ is invertible, and it maps the forward cone into itself because $U_{e-p_j}$ is positive; so $M_{ji} \in \mathbb{R}_{>0} \times O^+(1, 9)$, a Lorentz similitude.

[182] McCrimmon, *A Taste of Jordan Algebras*, note 168. II.4.2.2(3).

**The orthogonal case.** If $o_{ij} = 0$, let $r = e - p_i - p_j$, the third atom of a frame through $p_i, p_j$. Then $J_0(p_i) = \mathbb{R}p_j \oplus \mathbb{R}r \oplus J_{jr}$, where $J_{jr}$ is the off-diagonal Peirce space of $p_j, r$, and $U_{e-p_j}$ kills $p_j$ and $J_{jr}$ and fixes $r$. So

$$T_{ji}y = \langle r, y\rangle\, r \qquad (y \in J_0(p_i)), \tag{D.5}$$

with $0 \le r \le e - p_i$.

**Orientation.** The spinor lift below needs $\det M_{ji} > 0$ whenever $o_{ij} > 0$, and the coordinates are chosen to ensure it. The compact group $F_4 = \mathrm{Aut}(J)$ is a connected Lie group acting transitively on the atoms, and the stabiliser of an atom is the connected group Spin(9), so the atoms form the manifold $\mathbb{O}P^2 = F_4/\mathrm{Spin}(9)$ of dimension $16$[183].

Fix an atom $p_\star$ with coordinates $\varphi_\star = \varphi_{p_\star}$, choose $g_i \in F_4$ with $g_i p_\star = p_i$, and put $\varphi_{p_i} = g_i \circ \varphi_\star$. Since $gU_a g^{-1} = U_{ga}$ for automorphisms,

$$M_{ji} = \varphi_\star^{-1}\, U_{e-p_\star}\, h\, \varphi_\star, \qquad h = g_j^{-1} g_i, \qquad o_{ij} = \langle p_\star, h p_\star\rangle.$$

The set $K_+ = \{h \in F_4 : \langle p_\star, hp_\star\rangle > 0\}$ is connected: it is the preimage, under $h \mapsto hp_\star$, of the complement in $\mathbb{O}P^2$ of the closed submanifold of atoms orthogonal to $p_\star$ (an 8-sphere), which is connected, and the fibres are connected. On $K_+$ the map $h \mapsto \det(\varphi_\star^{-1} U_{e-p_\star} h \varphi_\star)$ is continuous and nonzero, and it equals 1 at $h = 1$. So $\det M_{ji} > 0$ whenever $o_{ij} > 0$.

**The Clifford representation.** On $\mathbb{R}^{16} = \mathbb{O}^2$ let $\Psi_9 = \mathrm{diag}(I_8, -I_8)$ and

$$\Psi_k = \begin{pmatrix} 0 & L_{e_{k-1}} \\ L_{e_{k-1}}^{\mathsf T} & 0 \end{pmatrix} \quad (k = 1, \dots, 8),$$

with $L_a$ the left multiplication by $a$ and $e_0 = 1, e_1, \dots, e_7$ the octonion units. Each $\Psi_k$ is symmetric with $\Psi_k^2 = I$, since $L_a$ is orthogonal for a unit $a$, and they anticommute: for orthogonal units $a, b$, $L_a L_b^{\mathsf T} + L_b L_a^{\mathsf T} = L_a L_{\bar b} + L_b L_{\bar a} = 2\,\mathrm{Re}(a\bar b) = 0$ by alternativity, and likewise for $L^{\mathsf T}L$; $\Psi_9$, block-diagonal with blocks $\pm I_8$, anticommutes with the block off-diagonal $\Psi_1, \dots, \Psi_8$. So

$$\Psi(x) = x_0 I + \textstyle\sum_{k=1}^{9} x_k \Psi_k$$

is a Jordan homomorphism from the spin factor into the real symmetric $16 \times 16$ matrices, $\Psi(x)$ has eigenvalues $x_0 \pm |\mathbf{x}|$, each eight times, and $\mathrm{Tr}\,\Psi(x) = 16x_0$. For $a, x \in \mathbb{R}^{1,9}$,

$$\tfrac{1}{8}\,\mathrm{Tr}\big(\Psi(a)\Psi(x)\big) = 2(a_0 x_0 + \mathbf{a}\cdot\mathbf{x}) = \mathrm{tr}\big(\varphi_p(a) \circ \varphi_p(x)\big). \tag{D.6}$$

Encode $y \in J_0(p_i)$, $y \ge 0$, as $\mathrm{enc}_i(y) = \Psi(\varphi_{p_i}^{-1} y)/8$, a positive semidefinite matrix of trace $\mathrm{tr}\, y$.

[183] Baez, "The octonions", note 5. §4.2.

**The spinor lift.** Let $g \in \mathbb{R}_{>0} \times O^+(1,9)$ with $\det g > 0$. Put $c = g(1)^{1/2}$, the square root in the spin factor of the interior point $g(1)$. Then $U_c$ is a similitude with factor $q_M(c)^2 = q_M(g(1))$, which is the factor of $g$; so $k = U_c^{-1} g$ lies in $O(1,9)$, fixes 1, and has $\det k > 0$, that is, $k \in SO(9)$.

For a unit $v \in \mathbb{R}^9$ and $\mathbf{x} \in \mathbb{R}^9$, $\Psi(v)\Psi(\mathbf{x})\Psi(v) = -\Psi(R_v \mathbf{x})$ with $R_v$ the reflection in $v^\perp$, while $\Psi(v)\Psi(v) = I$; so $S = \Psi(v)\Psi(w)$ satisfies $S\Psi(x)S^{\mathsf{T}} = \Psi(R_v R_w x)$ for all $x \in \mathbb{R}^{1,9}$, $R_v$ acting on the vector part. Every element of $SO(9)$ is a product of an even number of reflections, so there is an orthogonal $S$ with $S\Psi(x)S^{\mathsf{T}} = \Psi(kx)$. Since $\Psi$ is a Jordan homomorphism into an associative algebra, $\Psi(c)\Psi(y)\Psi(c) = \Psi(U_c y)$. So $A = \Psi(c)S$ satisfies

$$A\,\Psi(x)\,A^{\mathsf{T}} = \Psi(gx), \qquad A^{\mathsf{T}}A = S^{\mathsf{T}}\Psi\big(g(1)\big)S. \tag{D.7}$$

For $g = M_{ji}$ with $o_{ij} > 0$ let $A_{ji}$ be this lift. Here $g(1) = \varphi_{p_j}^{-1} U_{e-p_j}(e - p_i)$. Since $U_{e-p_j} p_i$ is $(1-o_{ij})$ times an atom of $J_0(p_j)$ (Lemma B.2), $U_{e-p_j}(e-p_i) = (e-p_j) - U_{e-p_j} p_i$ has eigenvalues 1 and $o_{ij}$ in $J_0(p_j)$. So $A_{ji}^{\mathsf{T}} A_{ji}$ has eigenvalues 1 and $o_{ij}$, each eight times, and $I - A_{ji}^{\mathsf{T}} A_{ji} \geq 0$ has eigenvalues 0 and $1 - o_{ij}$.

**The space and the instruments.** Let

$$\mathbb{C}^{17n+1} = \mathbb{C}|\mathrm{init}\rangle \oplus \bigoplus_{k=1}^{n} \mathbb{C}|P_k\rangle \oplus \bigoplus_{i=1}^{n} \mathbb{C}_i^{16},$$

with $\sigma = |\mathrm{init}\rangle\langle\mathrm{init}|$. A weight $t$ on $|P_k\rangle$ encodes $t\,p_k$, and $\mathrm{enc}_i(y)$ in block $\mathbb{C}_i^{16}$ encodes $y \in J_0(p_i)_+$. Let $\Pi_b$ be the projection onto block $b$. Each map below is applied to $\Pi_b X \Pi_b$ and the results are summed over $b$; discarding the coherences between blocks is completely positive and trace preserving. For test $j$:

(i) *One-dimensional blocks.* For $b = \mathrm{init}$ put $x_b = \rho$; for $b = P_k$ put $x_b = p_k$. Then

$$\mathcal{E}_{\mathrm{no}_j}(X) = \langle b|X|b\rangle\, \mathrm{enc}_j\big(U_{e-p_j} x_b\big), \qquad \mathcal{E}_{\mathrm{yes}_j}(X) = \langle b|X|b\rangle\, \langle p_j, x_b\rangle\, |P_j\rangle\langle P_j|.$$

(ii) *Block $j$.* $\mathcal{E}_{\mathrm{no}_j} = \mathrm{id}$ and $\mathcal{E}_{\mathrm{yes}_j} = 0$.

(iii) *Block $i \neq j$ with $o_{ij} > 0$.* $\mathcal{E}_{\mathrm{no}_j}(X) = A_{ji} X A_{ji}^{\mathsf{T}}$, placed in block $j$, and $\mathcal{E}_{\mathrm{yes}_j}(X) = \mathrm{Tr}\big((I - A_{ji}^{\mathsf{T}} A_{ji})X\big)\, |P_j\rangle\langle P_j|$.

(iv) *Block $i \neq j$ with $o_{ij} = 0$.* With $r$ as in (D.5) and $\Theta_i = \Psi(\varphi_{p_i}^{-1} r)$,

$$\mathcal{E}_{\mathrm{no}_j}(X) = \mathrm{Tr}(\Theta_i X)\, \mathrm{enc}_j(r), \qquad \mathcal{E}_{\mathrm{yes}_j}(X) = \mathrm{Tr}\big((I - \Theta_i)X\big)\, |P_j\rangle\langle P_j|.$$

*Complete positivity and normalisation.* In (i) each map is a positive functional times a positive matrix, and by (D.1) the traces add up to $\langle b|X|b\rangle$. In (ii) this is clear. In (iii) $I - A_{ji}^{\mathsf{T}} A_{ji} \geq 0$, and the traces add up to $\mathrm{Tr}\, X$. In (iv) $\Theta_i \geq 0$ because $r \geq 0$, and $I - \Theta_i = \Psi(\varphi_{p_i}^{-1}(e - p_i - r)) = \Psi(\varphi_{p_i}^{-1} p_j) \geq 0$; also $\mathrm{enc}_j(r)$ has trace one. So each pair is completely positive with trace-preserving sum, and all matrices are real.

**Induction.** Let $U_{(w)}$ denote the composite of the updates $U_{q_t}$ along a word $w$. *Claim:* $\mathcal{E}_w\sigma$ is the encoding of $U_{(w)}\rho$ in the block of the last letter of $w$ (the initial block if $w$ is empty): after $\mathrm{no}_j$, $U_{(w)}\rho \in J_0(p_j)_+$ and the state is $\mathrm{enc}_j(U_{(w)}\rho)$; after $\mathrm{yes}_j$, $U_{(w)}\rho = t\,p_j$ and the state is $t|P_j\rangle\langle P_j|$.

Suppose the claim for $w$ and apply test $j$. From a one-dimensional block holding $t\,x_b$, case (i) gives the encodings of $U_{e-p_j}(t x_b)$ and of $U_{p_j}(t x_b) = t\langle p_j, x_b\rangle p_j$. From block $j$, case (ii) is exact, since $U_{e-p_j}$ is the identity and $U_{p_j}$ is zero on $J_0(p_j)$.

From block $i \neq j$, let $y = U_{(w)}\rho = \varphi_{p_i}(x)$. In case (iii), by (D.7),

$$A_{ji}\,\tfrac{1}{8}\Psi(x)\,A_{ji}^{\mathsf{T}} = \tfrac{1}{8}\Psi(M_{ji}x) = \mathrm{enc}_j(T_{ji}y),$$

and the weight sent to $|P_j\rangle$ is $\operatorname{tr} y - \operatorname{tr} T_{ji}y = \langle p_j, y\rangle$ by (D.1), which encodes $U_{p_j}y = \langle p_j, y\rangle\, p_j$.

In case (iv), (D.6) gives $\mathrm{Tr}(\Theta_i\Psi(x)/8) = \langle r, y\rangle$, so the "no" output encodes $\langle r, y\rangle\, r = T_{ji}y$ by (D.5), and the "yes" weight is again $\operatorname{tr} y - \operatorname{tr} T_{ji}y$. This proves the claim. Taking traces, $\mathrm{Tr}(\mathcal{E}_w\sigma) = \mathrm{tr}(U_{(w)}\rho) = \Pr_\rho(w)$. □

# E Proper subalgebras and proofs for Section 7

This appendix proves Lemma 7.1, (7.3), Theorem 7.2(iii) and Proposition 7.4, in the notation of Section 7. The arguments below do not use the four-point invariant.

*Proof of Lemma 7.1.* For an idempotent $g \in B$ the Peirce projections (Section 3.1) are polynomials in the multiplication operator $x \mapsto g \circ x$, so they map $B$ into $B$. The component of $x \in H_3(\mathbb{K})$ in $F_{ab}(\mathbb{K})$ is obtained by applying the Peirce-$\frac{1}{2}$ projections of $E_{aa}$ and $E_{bb}$; so, with $S_{ab} = \{x : F_{ab}(x) \in B\}$, every off-diagonal entry of an element of $B$ in place $(a, b)$ lies in $S_{ab}$. Each $S_{ab}$ is a real subspace, $S_{ba} = \overline{S_{ab}}$, and by (7.1)

$$S_{ab}\,S_{bc} \subseteq S_{ac} \qquad (a, b, c \text{ distinct}).$$

Take $c = 3$; the other cases follow by permuting indices, which is conjugation by a real permutation matrix and hence an automorphism. From $1 \in S_{13} \cap S_{23}$ it follows, in turn,

$$\begin{aligned}
&1 = 1 \cdot 1 \in S_{13}S_{32} \subseteq S_{12},\\
&S_{13} = S_{13} \cdot 1 \subseteq S_{12}, \qquad S_{12} = S_{12} \cdot 1 \subseteq S_{13},\\
&S_{13} = 1 \cdot S_{13} \subseteq S_{21}S_{13} \subseteq S_{23}, \qquad S_{23} \subseteq S_{12}S_{23} \subseteq S_{13}.
\end{aligned}$$

So all three equal $\mathbb{F}$. Next $\overline{\mathbb{F}} = S_{21} \cdot 1 \subseteq S_{21}S_{13} \subseteq S_{23} = \mathbb{F}$ and $\mathbb{F}\mathbb{F} = S_{12}S_{23} \subseteq S_{13} = \mathbb{F}$.

The diagonal of a Hermitian matrix is real, so $B \subseteq H_3(\mathbb{F})$, and $H_3(\mathbb{F}) \subseteq B$ because $B$ contains the frame and all $F_{ab}(\mathbb{F})$. For the last claim, let $\mathbb{F}_0$ be the smallest such subalgebra. It lies in $\mathbb{F}$ by what was just shown. Conversely $H_3(\mathbb{F}_0)$ is closed under $\circ$ (entries of products are sums of products of entries and their conjugates) and contains $e$ and $X$, so $B \subseteq H_3(\mathbb{F}_0)$. □

The line subalgebra $\mathbb{R}q \oplus J_0(q)$ of a primitive idempotent $q$ (Section 7.3) is a unital subalgebra because $J_1(q) = \mathbb{R}q$, $J_0(q)$ is a subalgebra and $J_1 \circ J_0 = 0$. Its primitive idempotents are $q$ and the atoms on the line $e - q$.

**Lemma E.1** (Proper subalgebras)**.** *Let* $(\mathbb{K}, \mathbb{K}')$ *be* $(\mathbb{O}, \mathbb{H})$, $(\mathbb{H}, \mathbb{C})$ *or* $(\mathbb{C}, \mathbb{R})$*, and let* $B$ *be a proper unital subalgebra of* $H_3(\mathbb{K})$*. Then* $B$ *lies in a line subalgebra or in a unital subalgebra isomorphic to* $H_3(\mathbb{K}')$*. Every proper unital subalgebra of* $H_3(\mathbb{R})$ *lies in a line subalgebra.*

*Formalized.* `lem_five_sub`.

*Proof. Rank and summands.* $B$ is a Euclidean Jordan algebra with unit $e$. The trace of an idempotent of $H_3(\mathbb{K})$ is its rank, an integer. A Jordan frame of $B$ consists of orthogonal nonzero idempotents of $H_3(\mathbb{K})$ with sum $e$, so $B$ has rank at most three. If $B$ is not simple, its central idempotents have positive integer traces adding to 3, so one of them, $q$, has trace one and is primitive in $H_3(\mathbb{K})$. Since $q$ is central in $B$, every $x \in B$ equals $U_q x + U_{e-q} x \in \mathbb{R}q \oplus J_0(q)$.

*Simple of rank two.* By the classification (Section 3.1), $B$ is a spin factor with unit $e$. Its two orthogonal minimal idempotents have traces 1 and 2; let $c$ be the one of trace one. Pick $v = 2c' - e \in B$ with $c'$ an idempotent and $(2c - e) \circ v = 0$ (an element of the vector part orthogonal to $2c - e$). Expanding, $4\,c \circ c' - 2c - 2c' + e = 0$. Applying $U_c$, and using $U_c(c \circ x) = U_c x = \langle c, x\rangle c$, gives $(2\langle c, c'\rangle - 1)c = 0$. Taking the trace gives $4\langle c, c'\rangle + 1 - 2\,\mathrm{tr}\, c' = 0$, hence $\mathrm{tr}\, c' = \frac{3}{2}$, which is not an integer. So this case does not occur.

*Simple of rank one or three.* Rank one means $B = \mathbb{R}e$, which lies in every line subalgebra. Rank three means $B \cong H_3(\mathbb{K}'')$ with $\mathbb{K}'' \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}$ and $\dim B < \dim H_3(\mathbb{K})$; its primitive idempotents have trace one. If $\mathbb{K}'' = \mathbb{K}'$ the proof is complete. Two cases remain.

*A copy of* $H_3(\mathbb{R})$ *in* $H_3(\mathbb{H})$ *or* $H_3(\mathbb{C})$*.* Conjugating by a unitary matrix (an automorphism, as $\mathbb{K}$ is associative), move its frame to $E_{11}, E_{22}, E_{33}$[184]. Its Peirce spaces are lines $\mathbb{R}F_{ab}(x_{ab})$ with $|x_{ab}| = 1$, and (7.1) forces $x_{13} = \pm x_{12}x_{23}$. The diagonal unitary $\mathrm{diag}(1, x_{12}, x_{13})$ maps $F_{ab}(x)$ to $F_{ab}(d_a x \bar{d}_b)$, where $d_a$ is its $a$th diagonal entry, and makes all three entries real. So $B$ is conjugate to the standard $H_3(\mathbb{R}) \subset H_3(\mathbb{C})$.

*A copy of* $H_3(\mathbb{R})$ *or* $H_3(\mathbb{C})$ *in* $J$*.* Choose a unital isomorphism $\psi : H_3(\mathbb{K}'') \to B$, with $\mathbb{K}'' = \mathbb{R}$ or $\mathbb{C}$. The images of the standard frame are three nonzero orthogonal idempotents whose traces are positive integers summing to three; hence they form a frame of $J$. Move it to $(E_{11}, E_{22}, E_{33})$ by an automorphism. Peirce spaces then place $\psi(F_{ab}(1))$ in the slot $ab$, say $F_{ab}(u_{ab})$. Since $F_{ab}(1)^2 = E_{aa} + E_{bb}$, product preservation gives $|u_{ab}| = 1$.

By (F2) of Appendix A.5, a frame stabiliser makes $u_{12} = 1$. By (F3), another frame stabiliser, fixing $F_{12}(1)$, makes $u_{23} = 1$. The source relation $F_{12}(1) \circ F_{23}(1) = \frac{1}{2}F_{13}(1)$ then forces $\psi(F_{13}(1)) = F_{13}(1)$. Lemma 7.1 now gives $B = H_3(\mathbb{F})$ in these coordinates. Since $\dim B = 3 + 3\dim \mathbb{F}$ is 6 or 9, the algebra $\mathbb{F}$ has dimension one or two. It is therefore $\mathbb{R}$ or a complex plane in $\mathbb{O}$, and lies in a quaternion subalgebra by (F5). Undoing the automorphisms puts the original $B$ inside a unital copy of $H_3(\mathbb{H})$. □

For $J$ the lemma also follows from Racine's classification of maximal subalgebras of Albert algebras[185], specialised to the compact real form.

[184] Faraut and Korányi, *Analysis on Symmetric Cones*, note 16. Thm. IV.2.5.

[185] Racine, "Maximal subalgebras of exceptional Jordan algebras", note 172.

*Proof of* (7.3). $J_0(q) \cong H_2(\mathbb{K})$ is a spin factor, $a = \frac{1}{2}(f+u)$ and $b = \frac{1}{2}(f+u')$ with $u, u' \in J_0(q)$ orthogonal to $f$, and traceless elements of $H_2(\mathbb{K})$ satisfy $u \circ u' = \frac{1}{2}\langle u, u'\rangle f$. Hence $\operatorname{tr}(a \circ b) = \frac{1}{4}(2 + \langle u, u'\rangle) < 1$ for $a \neq b$, and

$$a + b - 2\,a \circ b = f + \tfrac{1}{2}(u+u') - \tfrac{1}{2}\big(f + u + u' + \tfrac{1}{2}\langle u, u'\rangle f\big) = \big(1 - \operatorname{tr}(a \circ b)\big) f.$$

□

**Lemma E.2** (Irredundant families)**.** *Let $S$ be an irredundant family of primitive idempotents of $H_3(\mathbb{K})$.*

*Formalized.* `lem_fiv_facts_a`, `lem_fiv_facts_b`.

*(a) (Matching.) No member of $S$ is orthogonal to two other members.*

*(b) If $S$ has at least three members and lies in the line subalgebra of $q$, then $q \notin S$ and all members of $S$ lie on the line $e - q$.*

*Proof.* (a) If $c \perp a$ and $c \perp b$ with $a \neq b$, then $a, b$ lie on the line $e - c$, and (7.3) recovers $e - c$, hence $c$, from $a$, $b$ and $e$. (b) The members other than $q$ are on the line $e - q$, hence orthogonal to $q$; by (a), $q \notin S$. □

**Lemma E.3** (Atoms on a line and one atom off it)**.** *Let $q$ be a primitive idempotent of $H_3(\mathbb{K})$, $f = e - q$, $p$ a primitive idempotent not in the line subalgebra of $q$, and $r = U_f p / \operatorname{tr} U_f p$. Let $s_1, \ldots, s_n$ ($n \geq 2$) be distinct atoms on the line $f$, none equal to $r$ or $f - r$. Then there is an automorphism $\alpha$ of $H_3(\mathbb{K})$ with*

*Formalized.* `lem_five_line`.

$$\alpha(q) = E_{22}, \quad \alpha(r) = E_{33}, \quad \alpha(f-r) = E_{11},$$
$$\alpha(p) = aE_{22} + (1-a)E_{33} + \sqrt{a(1-a)}\,F_{23}(1),$$

*where $0 < a < 1$, and $\alpha(s_k) = \gamma_k E_{11} + (1-\gamma_k)E_{33} + F_{13}(x_k)$ with $x_k \neq 0$ and $x_1 > 0$ real. Moreover* $\operatorname{Jord}\langle e, p, s_1, \ldots, s_n\rangle = \alpha^{-1}H_3(\mathbb{F})$*, where $\mathbb{F}$ is the smallest unital subalgebra of $\mathbb{K}$ closed under conjugation that contains $x_2, \ldots, x_n$.*

*Proof. The frame.* Put $a = \langle p, q\rangle$. If $a = 0$ then $p \in J_0(q)$ by Lemma 3.1, and if $a = 1$ then $p = q$; both put $p$ in the line subalgebra, so $0 < a < 1$. By Lemma B.2, $U_f p$ is a positive multiple of a primitive idempotent, of trace $\langle f, p\rangle = 1 - a$; so $r$ is an atom on the line $f$, and $(f - r, q, r)$ is a Jordan frame. Choose an automorphism $\alpha_0$ mapping it to $(E_{11}, E_{22}, E_{33})$[186].

*The shape of $p$ and of the $s_k$.* $\langle f - r, p\rangle = \langle f - r, U_f p\rangle = (1-a)\langle f - r, r\rangle = 0$, so $\alpha_0(p)$ lies in $J_0(E_{11})$ (Lemma 3.1), with diagonal entries $a$ at $E_{22}$ and $1 - a$ at $E_{33}$ and an entry $F_{23}(z)$. Here $z \neq 0$, since otherwise $\alpha_0(p) = aE_{22} + (1-a)E_{33}$ would not be an idempotent; and $|z|^2 = a(1-a)$, because a primitive idempotent of $J_0(E_{11}) \cong H_2(\mathbb{K})$ has determinant zero. Each $s_k \leq f$, so $\alpha_0(s_k)$ lies in $J_0(E_{22})$ and has an entry $F_{13}(x_k')$, with $x_k' \neq 0$ because $s_k \notin \{r, f - r\}$.

*Normalisation.* An automorphism is needed fixing $E_{11}, E_{22}, E_{33}$ that makes $z$ and $x_1'$ real and positive. For associative $\mathbb{K}$ the diagonal unitary $\operatorname{diag}(\bar{x}_1'/|x_1'|,\ \bar{z}/|z|,\ 1)$ does it, since it maps $F_{bc}(x)$ to $F_{bc}(d_b x \bar{d}_c)$, where $d_b$ is its $b$th diagonal entry.

[186] Faraut and Korányi, *Analysis on Symmetric Cones*, note 16. Thm. IV.2.5.

For $\mathbb{K} = \mathbb{O}$ the group Spin(8) acts on $J$ by automorphisms fixing the diagonal, on the entry in place $(2,3)$ by the vector representation and on the entry in place $(3,1)$ by a half-spin representation[187]. It is transitive on the unit sphere of the first, and the stabiliser Spin(7) of a unit vector there is transitive on the unit sphere of the second[188] (see also[189]). Let $\alpha$ be $\alpha_0$ followed by this automorphism.

*The generated algebra.* Let $B = \mathrm{Jord}\langle e, p, s_1, \dots, s_n\rangle$. It contains $f$ by (7.3) applied to $s_1, s_2$, hence $q$ and $U_f p$, hence $r$ and $f - r$. Peirce components of $\alpha(p)$ and $\alpha(s_1)$ give $F_{23}(1)$ and $F_{13}(1)$ in $\alpha(B)$, so Lemma 7.1 applies with $c = 3$. The off-diagonal entries of the generators are $\sqrt{a(1-a)}$, $x_1$ (both real) and $x_2, \dots, x_n$. □

*Proof of Theorem 7.2(iii).* The argument for (ii) in the proof of Theorem 7.2, without $p_4$, gives $H_3(\mathbb{H}_{ij})$, where $\mathbb{H}_{ij} = \mathrm{span}\{1, i, j, ij\}$ is the quaternion subalgebra generated by $i, j$; without $p_3$ it gives $H_3(\mathbb{H}_{i\ell})$. Both have dimension 15. Without $p_2$ the four atoms have entries in $\mathbb{H}_{j\ell}$, so they lie in $H_3(\mathbb{H}_{j\ell})$. They lie on the line $f = E_{11} + E_{22}$, so $f$ and $E_{33}$ lie in the generated algebra by (7.3). Writing $p_k = \frac{1}{2}(f + u_k)$, the four elements $u_k$ ($E_{11} - E_{22}$, $F_{12}(1)$, $F_{12}(-j)$, $F_{12}(-\ell)$) are linearly independent, and $u \circ u' = \frac{1}{2}\langle u, u'\rangle f$; so the generated algebra is $\mathbb{R}E_{33} \oplus \mathbb{R}f \oplus \mathrm{span}\{u_k\}$, of dimension 6.

For either subfamily without $p_0$ or without $p_1$, take $q = E_{33}$, $f = E_{11} + E_{22}$ and the off-line atom $p = p_2$. Its normalised line compression is

$$r = \frac{U_f p_2}{\operatorname{tr} U_f p_2} = p(1, i, 0), \qquad \operatorname{tr} U_f p_2 = \frac{2}{3}.$$

The other three atoms are distinct atoms on $f$, none equal to $r$ or $f - r$: their slot-12 entries are among $0, \frac{1}{2}, -\frac{j}{2}, -\frac{\ell}{2}$, whereas those of $r$, $f - r$ are $-\frac{i}{2}, \frac{i}{2}$. The slot 13 of $p_2$ is nonzero, so $p_2$ is not in the line subalgebra. Lemma E.3, with $n = 3$, therefore identifies the generated algebra with $H_3(\mathbb{F})$, where $\mathbb{F}$ is generated by two octonions and is contained in a quaternion subalgebra by Artin's theorem and (F5). Closure under conjugation adds no generator, since $\bar{x} = 2\operatorname{Re}(x) - x$.

It remains to see that $\mathbb{F}$ is quaternionic rather than real or complex. The generated algebra contains $f$ by (7.3) and contains $r$ by compression. Its corner at $f$ contains $f$ and four independent traceless elements:

$$\begin{aligned} &\text{without } p_0: \quad F_{12}(-i),\ F_{12}(1),\ F_{12}(-j),\ F_{12}(-\ell), \\ &\text{without } p_1: \quad F_{12}(-i),\ E_{11} - E_{22},\ F_{12}(-j),\ F_{12}(-\ell). \end{aligned}$$

These are $2r - f$ and $2s - f$ for the three retained line atoms; independence follows from the diagonal/slot decomposition and the orthogonal octonion basis. The corner thus has dimension at least five. Under the line lemma it is $H_2(\mathbb{F})$, of dimension $2 + \dim \mathbb{F}$. A unital real subalgebra of $\mathbb{H}$ is real, complex or quaternionic: two independent imaginary elements produce their independent cross product. Hence $\dim \mathbb{F} = 4$ and the generated algebra itself is a copy of $H_3(\mathbb{H})$, of dimension 15 in both cases. □

[187] Baez, "The octonions", note 5. §3.4.

[188] Ibid. §4.1.

[189] Kollross, *Octonions, triality, the exceptional Lie algebra $F_4$*, Table 3. proof of Thm. 3.1.

*Proof of Proposition 7.4. Lower bounds.* By Lemma 7.1, $E_{11}$, $p(1,1,0)$, $p(1,0,1)$ generate $H_3(\mathbb{R})$; any two primitive idempotents lie in a line subalgebra (take $q$ orthogonal to both), so the triple is irredundant.

The quadruple $E_{11}$, $E_{22}$, $p(1,1,1)$, $p(1,i,-1-i)$ generates $H_3(\mathbb{C})$ by the same lemma (the entry of $p(1,i,-1-i)$ in place $(1,2)$ is $-i/4$). Each of its triples lies in a copy of $H_3(\mathbb{R})$: $E_{11}, E_{22}, p(1,1,1)$ are real; for $E_{11}, E_{22}, p(1,i,-1-i)$ use the diagonal unitary $\mathrm{diag}(1,-i,(-1+i)/\sqrt{2})$; in the other two triples the vectors $(1,1,1)$ and $(1,i,-1-i)$ are orthogonal, and multiplying an orthonormal basis through them by phases makes the coordinates of the remaining vector real.

The subfamily $(p_0,p_1,p_2,p_3)$ of $\mathsf{F}$ generates $H_3(\mathbb{H}_{ij})$ (Theorem 7.2(iii)), and it is irredundant, since by Proposition 5.1 each of its triples lies in a complex subalgebra, of dimension $9 < 15$. Finally $\mathsf{F}$ gives 5 for $J$ (Theorem 7.2(ii),(iii)).

*Upper bounds.* Put $b(\mathbb{R},\mathbb{C},\mathbb{H},\mathbb{O}) = (3,4,4,5)$ and $\mathbb{K}'$ as in Lemma E.1. The following is proved by induction along $\mathbb{R},\mathbb{C},\mathbb{H},\mathbb{O}$:

*Claim.* Let $T$ be an irredundant family of primitive idempotents of $H_3(\mathbb{K})$ whose generated algebra $C$ lies in no line subalgebra. Then $|T| \le b(\mathbb{K})$. For $\mathbb{K} = \mathbb{C}$ and $|T| = 4$, moreover, every member of $T$ is orthogonal to exactly one other.

A generating family satisfies the hypothesis, so the Claim gives the upper bounds.

*Transport.* If $T$ lies in a unital subalgebra $Q$ with an isomorphism $\psi : H_3(\mathbb{K}'') \to Q$, then $\psi^{-1}T$ satisfies the hypothesis in $H_3(\mathbb{K}'')$, because $\psi$ maps a line subalgebra of $H_3(\mathbb{K}'')$ into a line subalgebra of $H_3(\mathbb{K})$ (it preserves Peirce spaces, and primitive idempotents of $Q$ have trace one).

Let $m = |T|$ and $C_t = \mathrm{Jord}\langle e, T \setminus \{t\}\rangle$, a proper subalgebra of $C$.

*Step 1: $C \neq H_3(\mathbb{K})$.* By Lemma E.1, $C$ lies in a copy of $H_3(\mathbb{K}')$, and transport with the Claim for $\mathbb{K}'$ gives $m \le b(\mathbb{K}') \le b(\mathbb{K})$. For $\mathbb{K} = \mathbb{R}$ this step cannot occur.

*Step 2: $C = H_3(\mathbb{K})$ and some $C_t$ lies in the line subalgebra of $q$.* Suppose $m \ge 4$. By Lemma E.2(b) the $n = m - 1 \ge 3$ members $s_1, \dots, s_n$ of $T \setminus \{t\}$ lie on the line $f = e - q$, and $p = t$ is not in the line subalgebra. No $s_k$ equals $r$ or $f - r$: these lie in the algebra generated by $p$ and two other $s$'s (as in the last paragraph of the proof of Lemma E.3), contradicting irredundancy. So Lemma E.3 applies.

Deleting $s_l$ with $l \ge 2$ leaves $p$, $s_1$ and at least one more $s$, so it generates $\alpha^{-1}H_3(\mathbb{F}_l)$, where $\mathbb{F}_l$ is generated by the $x_k$, $k \notin \{1,l\}$; irredundancy forces $\mathbb{F}_l \neq \mathbb{K}$.

- $\mathbb{K} = \mathbb{R}$: $\mathbb{F}_l = \mathbb{R} = \mathbb{K}$, a contradiction. So $m \le 3$.
- $\mathbb{K} = \mathbb{C}$: each $x_k$, $k \ge 2$, survives a deletion $l \ge 2$, so it is real. Then all $x_k$ are real and $T$ generates a copy of $H_3(\mathbb{R})$, a contradiction. So $m \le 3$.
- $\mathbb{K} = \mathbb{H}$, $m \ge 5$: each pair $x_a, x_b$ $(a,b \ge 2)$ survives a deletion, and a proper conjugation-closed subalgebra of $\mathbb{H}$ is commutative; commuting quaternions have parallel imaginary parts. So all $x_k$ lie in one copy of $\mathbb{C}$, a contradiction. So $m \le 4$.
- $\mathbb{K} = \mathbb{O}$, $m \ge 6$: each triple $x_a, x_b, x_c$ $(a,b,c \ge 2)$ survives a deletion, so it lies in a proper conjugation-closed subalgebra. If the imaginary parts of all $x_k$ are parallel, they lie in one copy of $\mathbb{C}$. Otherwise two of them, $u$ and $v$, generate

a quaternion subalgebra $\mathbb{H}_{uv}$, and for every other $x_c$ the proper subalgebra containing $u, v, x_c$ contains $\mathbb{H}_{uv}$, so equals it: a conjugation-closed subalgebra of $\mathbb{O}$ strictly containing a quaternion subalgebra $H$ is $\mathbb{O}$, since $\mathbb{O} = H \oplus H\ell'$ for a unit $\ell' \perp H$ and $h(h'\ell') = (h'h)\ell'$[190]. So all $x_k$ lie in $\mathbb{H}_{uv}$, a contradiction. So $m \le 5$.

*Step 3: $C = H_3(\mathbb{K})$ and no $C_t$ lies in a line subalgebra.* Then every $C_t$ lies in a copy of $H_3(\mathbb{K}')$ (for $\mathbb{K} = \mathbb{R}$ this cannot happen), and transport with the Claim for $\mathbb{K}'$ gives $m - 1 \le b(\mathbb{K}')$. This is $m \le 4$ for $\mathbb{C}$ and $m \le 5$ for $\mathbb{O}$.

For $\mathbb{H}$ it gives $m \le 5$; if $m = 5$, each $T \setminus \{t\}$ has four members and, by the Claim for $\mathbb{C}$, its orthogonality graph is a perfect matching. But by Lemma E.2(a) the orthogonality graph of $T$ is a matching on five vertices, so it has an isolated vertex $v$, and deleting some $t \neq v$ leaves $v$ isolated. So $m \le 4$.

*The matching statement for $\mathbb{C}$.* Let $m = 4$. Step 1 would give $m \le 3$ and Step 2 excludes $m \ge 4$, so Step 3 applies: each triple lies in a copy of $H_3(\mathbb{R})$, which is unitarily conjugate to the standard one (proof of Lemma E.1).

Suppose some $a \in T$ is orthogonal to no other member. Write the members as rank-one projections onto unit vectors $\psi_t \in \mathbb{C}^3$, with phases chosen so that $\langle \psi_a, \psi_t \rangle > 0$. The Bargmann invariant $\langle \psi_a, \psi_s \rangle \langle \psi_s, \psi_t \rangle \langle \psi_t, \psi_a \rangle$ of a triple in a conjugate of $H_3(\mathbb{R})$ is real, so every $\langle \psi_s, \psi_t \rangle$ is real. A real Gram matrix means that a unitary moves all four vectors into $\mathbb{R}^3$, so $T$ lies in a copy of $H_3(\mathbb{R})$, contradicting $C = H_3(\mathbb{C})$. Hence every member is orthogonal to another, and by Lemma E.2(a) to exactly one. □

# F An algebraic proof of the obstruction

Four idempotents of an associative algebra satisfying the corner relations obey a polynomial identity, without positivity or square roots; the tetrahedron identity of Section 5.4 is its Hilbert-space case. Its consequence for the Albert algebra also supplies the final contradiction in Section 4.3.

### Thirteen traces and three corner relations

The numbers $o_{ij}, R_{ijk}, t_r, \omega_r$ are those of Section 5.2, and $Q_{j;klm} = \langle p_j, \{p_k p_l p_m\} \rangle$ is the number (5.1), written with the *Jordan triple product* $\{xyz\} = (x \circ y) \circ z + (z \circ y) \circ x - (x \circ z) \circ y$; so $Q_{j;klm} = Q_{j;mlk}$.

For an atom $p$ the rank-one formula $U_p y = \langle p, y \rangle\, p$ turns these numbers into statements about compressions:

$$U_{p_a} p_b = o_{ab}\, p_a, \qquad U_{p_a}(p_b \circ p_c) = R_{abc}\, p_a, \qquad U_{p_a}\{p_b p_c p_d\} = Q_{a;bcd}\, p_a.$$

In an associative algebra with $x \circ y = \frac{1}{2}(xy + yx)$ one has $U_x y = xyx$ and $\{xyz\} = \frac{1}{2}(xyz + zyx)$. This suggests the following relations.

[190] Springer and Veldkamp, *Octonions, Jordan Algebras and Exceptional Groups*, note 167. Ch. 1.

**Definition F.1** (Corner relations)**.** Let $A$ be an associative real algebra, not necessarily unital, and $X_A, X_B, X_C, X_D \in A$ idempotents. Let real numbers $o_{ab} = o_{ba}$, $R_{abc}$ (unchanged under permutation of the labels) and $Q_{a;bcd}$ be given. The *corner relations* are, for distinct labels,

*Formalized.* `CornerRelations`.

$$\text{(R1)} \quad X_a X_b X_a = o_{ab}\, X_a, \tag{F.1}$$

$$\text{(R2)} \quad X_a (X_b X_c + X_c X_b) X_a = 2R_{abc}\, X_a, \tag{F.2}$$

$$\text{(R3)} \quad X_a (X_b X_c X_d + X_d X_c X_b) X_a = 2Q_{a;bcd}\, X_a. \tag{F.3}$$

Each relation says that an element of $A$, squeezed into the corner $X_a A X_a$, becomes a prescribed multiple of $X_a$. Only reversal-symmetric combinations are prescribed, never a single product such as $X_a X_b X_c X_a$.

**Lemma F.2** (Jordan homomorphisms give the corner relations)**.** *Let $C \subseteq J$ be a Jordan subalgebra (not necessarily unital) containing atoms $p_A, \dots, p_D$, let $A$ be an associative real algebra, and let $\eta : C \to A^+$ be a Jordan homomorphism (Section 3.1), so that $\eta(x \circ y) = \frac{1}{2}\big(\eta(x)\eta(y) + \eta(y)\eta(x)\big)$. Then*

*Narrower in Lean.* `jordanHom_map_quadratic`, `jordanHom_map_triple`. Only for maps defined on all of $J$; the corner relations for a subalgebra $C$ are not stated.

$$\eta(U_x y) = \eta(x)\eta(y)\eta(x), \qquad \eta(\{xyz\}) = \tfrac{1}{2}\big(\eta(x)\eta(y)\eta(z) + \eta(z)\eta(y)\eta(x)\big)$$

*for $x, y, z \in C$, and the elements $X_a = \eta(p_a)$ are idempotents satisfying (R1)–(R3) with the thirteen traces of $p_A, \dots, p_D$.*

*Proof.* Write $X = \eta(x)$, $Y = \eta(y)$, $Z = \eta(z)$. Then $\eta(x \circ x) = X^2$ and

$$\eta(U_x y) = \eta\big(2x\circ(x\circ y) - (x\circ x)\circ y\big) = \tfrac{1}{2}\big(X^2Y + 2XYX + YX^2\big) - \tfrac{1}{2}\big(X^2Y + YX^2\big) = XYX.$$

For the triple product, $\eta((x \circ y) \circ z) = \frac{1}{4}(XYZ + YXZ + ZXY + ZYX)$. Adding the same expression with $x, z$ exchanged and subtracting the one with $y, z$ exchanged leaves $\frac{1}{2}(XYZ + ZYX)$. In particular $X_a^2 = \eta(p_a \circ p_a) = X_a$. The elements $p_b$, $p_b \circ p_c$ and $\{p_b p_c p_d\}$ lie in $C$, so applying $\eta$ to the three compressions above gives

$$\begin{aligned} X_a X_b X_a &= o_{ab} X_a, \\ X_a \tfrac{1}{2}(X_b X_c + X_c X_b) X_a &= R_{abc} X_a, \\ X_a \tfrac{1}{2}(X_b X_c X_d + X_d X_c X_b) X_a &= Q_{a;bcd} X_a. \end{aligned}$$

□

From $J$ this proof uses only the rank-one formula and the definitions of the traces. The same computation applies to rank-one projections $P_a = a_a a_a^*$ in $M_3(\mathbb{C})$, with the traces taken in $H_3(\mathbb{C})$.

## The identity

Fix a base label $j$ and an order $k, l, m$ of the other three. The three triangles at $j$ carry the shape invariants and products of transition probabilities

$$r_1 = R_{jkl}, \quad r_2 = R_{jlm}, \quad r_3 = R_{jmk},$$
$$\kappa_1 = o_{jk}o_{kl}o_{lj}, \quad \kappa_2 = o_{jl}o_{lm}o_{mj}, \quad \kappa_3 = o_{jm}o_{mk}o_{kj}.$$

Let $\mathcal{B} = (b_{ab})$ be the symmetric $3 \times 3$ matrix with entries

$$b_{aa} = r_a^2 - \kappa_a \quad (a = 1, 2, 3),$$
$$b_{12} = o_{jl}\, Q_{j;klm} - r_1 r_2, \qquad b_{23} = o_{jm}\, Q_{j;lmk} - r_2 r_3, \qquad b_{13} = o_{jk}\, Q_{j;mkl} - r_3 r_1,$$
$$\sigma = o_{jk} o_{jl} o_{jm}\, R_{klm} - r_1 r_2 r_3 - r_1 b_{23} - r_2 b_{13} - r_3 b_{12}, \qquad \Delta_j = -\det \mathcal{B} - \sigma^2. \tag{F.4}$$

For four atoms $\Delta_j$ is a polynomial in the thirteen traces. It does not depend on the order of $k, l, m$: a cyclic shift of $(k, l, m)$ permutes the indices of $\mathcal{B}$ cyclically and fixes $\sigma$, and the reversal $(k, l, m) \mapsto (k, m, l)$ exchanges the indices 1 and 3 of $\mathcal{B}$ and fixes $\sigma$ (using $Q_{j;klm} = Q_{j;mlk}$). For positive transition probabilities $\Delta_j = (\nu_1 \nu_2 \nu_3)^2\, \mathcal{D}$, with $\nu_a = \sqrt{\kappa_a}$ the normalisations of Section 5.4 (proof of Theorem 5.9). The proof below does not use the four-point invariant or its symbolic expansion.

**Theorem F.3** (Tetrahedron identity)**.** *Let $A$ be an associative real algebra and $X_j, X_k, X_l, X_m \in A$ idempotents. Let real numbers $o_{ab} = o_{ba}$, $R_{abc}$ (symmetric in its labels) and $Q_{j;klm}$, $Q_{j;lmk}$, $Q_{j;mkl}$ be given such that (R1) holds for all ordered pairs of distinct labels, (R2) holds at the base $j$ for the pairs $\{k, l\}$, $\{l, m\}$, $\{m, k\}$ and at the base $k$ for the pair $\{l, m\}$, and (R3) holds at the base $j$ for the triples $(k, l, m)$, $(l, m, k)$, $(m, k, l)$. Define $\Delta_j$ by* (F.4)*. Then*

$$\Delta_j\, X_j = 0.$$

*Formalized.* `tetrahedronEngine`, with a possibly nonunital target.

No transition probability needs to be nonzero, and the trace relation of Section 5 is not used. Its precedent is discussed in Section 5.4.

*Proof.* Every element below lies in the corner $A_j = X_j A X_j$, an associative real algebra with unit $X_j$ (if $X_j = 0$ there is nothing to prove). For a real number $c$ write $c$ also for $c X_j$. Put

$$g_1 = X_j X_k X_l X_j, \quad g_2 = X_j X_l X_m X_j, \quad g_3 = X_j X_m X_k X_j,$$
$$g_1' = X_j X_l X_k X_j, \quad g_2' = X_j X_m X_l X_j, \quad g_3' = X_j X_k X_m X_j.$$

Each step uses (R1) only to contract a product $X_a X_b X_a$ inside a longer word.

*(a) Reversal.* (R2) at $j$ for the pair $\{k, l\}$ says $g_1 + g_1' = 2r_1$; likewise $g_a + g_a' = 2r_a$ for $a = 2, 3$. Put $n_a = g_a - r_a$. Then

$$g_a = r_a + n_a, \qquad g_a' = r_a - n_a.$$

*(b) Length.* Since $X_j^2 = X_j$, contracting $X_l X_j X_l = o_{jl} X_l$, then $X_k X_l X_k = o_{kl} X_k$, then $X_j X_k X_j = o_{jk} X_j$ gives

$$g_1 g_1' = X_j X_k (X_l X_j X_l) X_k X_j = o_{jl}\, X_j (X_k X_l X_k) X_j = o_{jl} o_{kl}\, X_j X_k X_j = \kappa_1,$$

and in the same way $g_2 g_2' = \kappa_2$, $g_3 g_3' = \kappa_3$.

*(c) Squares.* By (a) and (b), $\kappa_a = (r_a + n_a)(r_a - n_a) = r_a^2 - n_a^2$, so $n_a^2 = b_{aa}$.

*(d) Anticommutators.* Contracting $X_l X_j X_l = o_{jl} X_l$ in the middle,

$$g_1 g_2 = o_{jl}\, X_j X_k X_l X_m X_j, \qquad g_2' g_1' = o_{jl}\, X_j X_m X_l X_k X_j,$$

and (R3) at $j$ gives $g_1g_2+g_2'g_1' = 2o_{jl}Q_{j;klm}$. On the other hand, by (a), $g_1g_2+g_2'g_1' = (r_1+n_1)(r_2+n_2)+(r_2-n_2)(r_1-n_1) = 2r_1r_2+n_1n_2+n_2n_1$. The pairs $(g_2,g_3)$ and $(g_3,g_1)$ work the same way with the contractions $X_mX_jX_m = o_{jm}X_m$ and $X_kX_jX_k = o_{jk}X_k$, so

$$n_an_b+n_bn_a = 2b_{ab} \qquad (a,b=1,2,3). \tag{F.5}$$

*(e) The far face.* Contracting $X_lX_jX_l$ and then $X_mX_jX_m$,

$$g_1g_2g_3 = o_{jl}o_{jm}\,X_jX_kX_lX_mX_kX_j, \qquad g_3'g_2'g_1' = o_{jl}o_{jm}\,X_jX_kX_mX_lX_kX_j.$$

Adding, (R2) at the base $k$ for the pair $\{l,m\}$ and then $X_jX_kX_j = o_{jk}X_j$ give

$$g_1g_2g_3+g_3'g_2'g_1' = o_{jl}o_{jm}\,X_jX_k(X_lX_m+X_mX_l)X_kX_j = 2\,o_{jk}o_{jl}o_{jm}R_{klm}. \tag{F.6}$$

Expand the left side with (a). The terms of degree zero and two in the $n_a$ come in pairs:

$$\begin{aligned}(r_1+n_1)(r_2+n_2)(r_3+n_3)+(r_3-n_3)(r_2-n_2)(r_1-n_1)\\ = 2r_1r_2r_3+r_1(n_2n_3+n_3n_2)+r_2(n_1n_3+n_3n_1)\\ +r_3(n_1n_2+n_2n_1)+n_1n_2n_3-n_3n_2n_1.\end{aligned}$$

the terms of degree one cancelling. By (F.5) the three anticommutator terms add up to $2(r_1b_{23}+r_2b_{13}+r_3b_{12})$. Comparing with (F.6),

$$\Gamma := \tfrac{1}{2}\big(n_1n_2n_3-n_3n_2n_1\big) = \sigma. \tag{F.7}$$

*(f) The volume element.* By (F.5) the elements $N_a = n_a$ of the unital associative algebra $A_j$ satisfy $N_aN_b+N_bN_a = -2M_{ab}X_j$ with $M=-\mathcal{B}$. Lemma A.10, applied in $A_j$, gives $\Gamma^2 = \det M\,X_j = -\det\mathcal{B}\,X_j$, since $\mathcal{B}$ is $3\times3$. With (F.7), $\Gamma^2=\sigma^2X_j$, so $(-\det\mathcal{B}-\sigma^2)X_j=0$, that is, $\Delta_jX_j=0$. □

## A configuration that violates it

Let $\ell = e_4$ be the octonion unit of Appendix A.2 and put

$$\begin{gathered}v_A=(1,0,0),\quad v_B=\tfrac{1}{\sqrt2}(1,1,0),\quad v_C=\tfrac{1}{\sqrt3}(1,i,1),\\ v_D=\tfrac{1}{\sqrt3}(1,j,\ell),\qquad s_a=v_av_a^\dagger.\end{gathered} \tag{F.8}$$

The entries of each $v_a$ lie in an associative subalgebra of $\mathbb{O}$, so each $s_a$ is an atom (Appendix A.4); the configuration $(s_A,s_B,s_C,s_D)$ is the configuration $\mathsf{S}$ of Section 6. Explicitly, in the notation (A.2) with entries $(\xi_1,\xi_2,\xi_3;x_1,x_2,x_3)$,

$$\begin{aligned}s_A=(1,0,0;0,0,0),\quad s_B=(\tfrac12,\tfrac12,0;0,0,\tfrac12),\\ s_C=\tfrac13(1,1,1;i,1,-i),\quad s_D=\tfrac13(1,1,1;-j\ell,\ell,-j).\end{aligned}$$

Its traces are listed in Section 6 (paragraph "Two configurations") and derived in the proof of Proposition 6.11; in particular all three $\omega_r$ and all twelve $Q$'s equal $\frac{1}{18}$.

**The values of $\Delta_j$.** At the base $A$ with order $(B, C, D)$: $r = (\frac{1}{6}, \frac{1}{9}, \frac{1}{6})$, $\kappa = (\frac{1}{18}, \frac{1}{27}, \frac{1}{18})$, $b_{11} = b_{33} = \frac{1}{36} - \frac{1}{18} = -\frac{1}{36}$, $b_{22} = \frac{1}{81} - \frac{1}{27} = -\frac{2}{81}$, and every off-diagonal entry vanishes, for example $b_{13} = o_{AB}Q_{A;DBC} - r_3 r_1 = \frac{1}{36} - \frac{1}{36} = 0$. So $\det \mathcal{B} = -\frac{2}{104976}$, $\sigma = \frac{1}{2}\cdot\frac{1}{3}\cdot\frac{1}{3}\cdot\frac{1}{9} - \frac{1}{324} = \frac{1}{324}$ and

$$\Delta_A(\mathsf{S}) = \tfrac{2}{104976} - \tfrac{1}{104976} = \tfrac{1}{104976}.$$

The base $B$ with order $(A, C, D)$ gives the same $r$, $\kappa$, $\mathcal{B}$ and $\sigma$, so $\Delta_B(\mathsf{S}) = \frac{1}{104976}$. At the base $C$ with order $(A, B, D)$: $r = (\frac{1}{6}, \frac{1}{9}, \frac{1}{9})$, $\kappa = (\frac{1}{18}, \frac{1}{27}, \frac{1}{27})$, $\mathcal{B}$ has diagonal $(-\frac{1}{36}, -\frac{2}{81}, -\frac{2}{81})$ and the single nonzero off-diagonal pair $b_{23} = \frac{1}{54} - \frac{1}{81} = \frac{1}{162}$, so $\det \mathcal{B} = -\frac{15}{944784}$, $\sigma = \frac{1}{162} - \frac{1}{486} - \frac{1}{972} = \frac{1}{324}$ and $\Delta_C(\mathsf{S}) = \frac{6}{944784} = \frac{1}{157464}$; the base $D$ is the same.

**A cubic equation.** Let $y = s_A + s_B + s_C + s_D$. Every $x \in J$ satisfies $x^3 - T(x)x^2 + S(x)x - N(x)e = 0$, with $x^3 = x \circ x^2$, $T = \mathrm{tr}$, $S(x) = \frac{1}{2}(T(x)^2 - \langle x, x\rangle)$ and the cubic norm $N$ (the cubic minimal equation, Appendix A.3; McCrimmon, Construction 4.2.2, p. 190[191]). Here $T(y) = 4$, $\langle y, y\rangle = 4 + 2\sum o_{ab} = \frac{25}{3}$, so $S(y) = \frac{23}{6}$, and $\mathrm{tr}(y^3) = 4 + 6\sum o_{ab} + 6\sum R_{abc} = \frac{61}{3}$, by expanding the trace and using its associativity. The cubic Newton identity gives

$$N(y) = \tfrac{1}{6}\big(T(y)^3 - 3T(y)\,\mathrm{tr}(y^2) + 2\,\mathrm{tr}(y^3)\big) = \tfrac{1}{6}(64 - 100 + 122/3) = \tfrac{7}{9}.$$

Hence

$$y^3 - 4y^2 + \tfrac{23}{6}\,y - \tfrac{7}{9}\,e = 0, \qquad e = \tfrac{9}{7}\big(y^3 - 4y^2 + \tfrac{23}{6}\,y\big).$$

## Two consequences

**Corollary F.4** (No special subalgebra contains them)**.** *Let $p_A, p_B, p_C, p_D$ be atoms of $J$ with $\Delta_j \neq 0$ for some base $j$. If $C \subseteq J$ is a Jordan subalgebra containing them and $\eta : C \to A^+$ is a Jordan homomorphism, with $A$ an associative real algebra, then $\eta(p_j) = 0$. In particular no Jordan subalgebra of $J$ containing the four atoms admits an injective Jordan homomorphism into $A^+$ for any associative $A$: the four atoms lie in no special subalgebra of $J$.*

*Narrower in Lean.* `albertJordanHom_eq_zero`, `proper_iff_special`. Only the case $C = J$, where $\eta = 0$.

*Proof.* By Lemma F.2 the $\eta(p_a)$ satisfy (R1)–(R3), and Theorem F.3 gives $\Delta_j \eta(p_j) = 0$, so $\eta(p_j) = 0$. An injective $\eta$ would then give $p_j = 0$, but atoms are nonzero. □

**Theorem F.5** (The Albert algebra is exceptional)**.** *Let $A$ be an associative real algebra and $\eta : J \to A^+$ a Jordan homomorphism. Then $\eta = 0$. In particular $J$ is not special, and a unital Jordan homomorphism from $J$ into $A^+$ exists only for $A = 0$.*

*Different route.* `albertJordanHom_eq_zero`: the unit is reached by Peirce propagation from S instead of the cubic equation.

*Proof.* Put $X_a = \eta(s_a)$ for the configuration S. Corollary F.4 at the base $A$, where $\Delta_A(\mathsf{S}) = \frac{1}{104976} \neq 0$, gives $X_A = 0$. For each $b \neq A$, (R1) at the base $b$ reads $X_b X_A X_b = o_{Ab} X_b$, and $o_{Ab} \in \{\frac{1}{2}, \frac{1}{3}\}$ is nonzero, so $X_b = 0$. Hence $\eta(y) = 0$ for $y = \sum_a s_a$, then $\eta(y^2) = \eta(y)^2 = 0$ and $\eta(y^3) = \frac{1}{2}(\eta(y)\eta(y^2) + \eta(y^2)\eta(y)) = 0$. The cubic equation for $y$ gives $\eta(e) = \frac{9}{7}\eta(y^3 - 4y^2 + \frac{23}{6}y) = 0$. Finally $U_e x = x$ for every $x$, so $\eta(x) = \eta(U_e x) = \eta(e)\eta(x)\eta(e) = 0$. □

[191] McCrimmon, *A Taste of Jordan Algebras*, note 168.

Albert proved in 1934 that $J$ is not special[192]. One standard proof exhibits an identity that holds in every special Jordan algebra but fails in $J$, such as Glennie's identity[193]. The proof above uses neither an identity of that kind nor the simplicity of $J$: the obstruction is the single number $\Delta_A(\mathsf{S})$, together with one cubic equation that brings back the unit. Theorem F.3 concerns relations among particular elements, not an identity in free variables, so it does not bear on the minimal degree of identities separating special from general Jordan algebras.

### Glennie's identity at four atoms

For $x, y$ in a Jordan algebra put $D_{x,y} = 4[L_x, L_y]$, $U_{[x,y]} = 4U_{x\circ y} - 2(U_xU_y + U_yU_x)$, $D^{\mathrm{C}}_{x,y} = D^3_{x,y} + 3D_{x,y}U_{[x,y]}$,

$$\psi(x,y,z) = D^{\mathrm{C}}_{x,y}(z^2) - 2z\circ D^{\mathrm{C}}_{x,y}(z), \qquad F(a,b,c,d) = \left[s^2t^2u^2\right]\psi(sa+tb,\ sa+uc,\ d).$$

This is Shestakov's form of Glennie's identity of degree eight[194], as recorded by McCrimmon[195] (there written with $V_x = 2L_x$ and $\{x,y\} = 2x\circ y$; McCrimmon states $\psi = 6G_8$, which is not used here). In $A^+$, $A$ associative, with $q = xy - yx$ one has $D_{x,y}z = qz - zq$, $U_{[x,y]}z = qzq$ and $D^{\mathrm{C}}_{x,y}z = q^3z - zq^3$, a derivation; so $\psi$ and its polarisation $F$, homogeneous of degree two in each of four variables, vanish in every special Jordan algebra.

**Theorem F.6** (The polarised Glennie identity detects $\mathcal{A}$)**.** *Let $p_A, p_B, p_C, p_D$ be atoms of $J$ and $r_x = 2p_x - e$. Then*

$$\|F(r_A, r_B, r_C, r_D)\|^2 = 9\cdot 2^{21}\, N(p_A+p_B+p_C)\,\mathcal{A}(p_A,p_B,p_C,p_D),$$

*and the same holds with any three of the atoms, in any order, in the first three places and $N$ of their sum. In particular $F$ vanishes at the four symmetries if and only if $\mathcal{A} = 0$.*

*Different form in Lean.* `AppF.thm_glennie`, `AppF.thm_glennie_normSq`, `AppF.thm_glennie_perm`, `AppF.thm_glennie_eq_zero_iff`. Lean defines $F$ as the sum of the nine placements of the polarisation of $\psi$; that this is the coefficient $[s^2t^2u^2]$ is not a Lean theorem. The values for M and S after the proof: `AppF.glennie_maximizer_value`, `AppF.glennie_fixedS_value`, `AppF.glennie_fixedS_ABD_C`, `AppF.glennie_fixedS_ACD_B`, `AppF.glennie_fixedS_BCD_A`.

*Proof.* If $\Delta_W(p_A,p_B,p_C) = 0$, then $\mathcal{A} = 0$ (Proposition 5.6) and the four atoms lie in a quaternionic subalgebra (Theorem 5.5(ii)), which is special, so $F = 0$. Otherwise, by Lemma A.9(b) and $F_4$-invariance ($\psi$ is a Jordan polynomial and $F_4$ acts by automorphisms orthogonal for the trace form, so $\|F\|^2$, $N$ and $\mathcal{A}$ are invariant), $p_A, p_B, p_C \in W = H_3(\mathbb{C}_i)$; put $N_W = N(p_A+p_B+p_C)$ and $\beta = \mathrm{Tr}(p_Ap_Bp_C)$ (complex matrices), so $\Delta_W = N_W(\mathrm{Im}\,\beta)^2$.

*The complex subalgebra acts associatively.* $M_3(\mathbb{O})$ is an associative $M_3(\mathbb{C}_i)$-bimodule, since an associator with two arguments in $\mathbb{C}_i$ vanishes. So for $x, y \in W$ the formulas above for $D_{x,y}$, $U_{[x,y]}$, $D^{\mathrm{C}}_{x,y}$ hold as operators on $J$, and extracting $s^2t^2u^2$ from $q^3$, $q = -st[r_A,r_B] + su[r_A,r_C] + tu[r_B,r_C]$, gives $F(r_A,r_B,r_C,z) = Yz^2 - z^2Y - 2z\circ(Yz - zY)$, where $Y$ is the negative of the sum of the six products of the three commutators. $Y$ is skew-Hermitian; $Y - i\lambda 1 \in \mathfrak{su}(3)$, $\lambda = \mathrm{Tr}\,Y/3i$, is a sum of commutators $[h,h']$ of Hermitian complex matrices, and $z \mapsto [h,h']z - z[h,h']$ is the inner derivation

[192] Albert, "On a certain algebra of quantum mechanics", note 6.
[193] Glennie, "Some identities valid in special Jordan algebras", note 7.
[194] Ibid.
[195] McCrimmon, "The role of identities in Jordan algebras", note 60. §§11, 13, pp. 276–277.

$D_{h,h'}$ of $J$[196], which drops out. With $B(z) = iz - zi$ (entrywise), $z = r_D$, $z^2 = e$ and $B(e) = 0$:

$$F(r_A, r_B, r_C, r_D) = -2\lambda\, r_D \circ B(r_D) = 4\lambda\,\mathrm{Def}, \qquad \mathrm{Def} = B(p_D) - 2p_D \circ B(p_D).$$

*The constant.* Expanding the three commutators of rank-one projections, every trace of a product of six of them is a product of transition probabilities and of $\beta, \bar{\beta}$; one finds $\mathrm{Tr}\, Y = 384\, i\, N_W \,\mathrm{Im}\,\beta$, so $\lambda = 128\, N_W \,\mathrm{Im}\,\beta$ and $\lambda^2 = 2^{14} N_W \Delta_W$.

*The defect.* For octonions, $B(uv) - B(u)v - uB(v) = -3[i,u,v]$, so $\mathrm{Def}_{ac} = -3\sum_b [i, (p_D)_{ab}, (p_D)_{bc}]$, and only the $\mathbb{C}_i^{\perp}$-components $m_{ab}$ of the off-diagonal entries contribute. For $u, v \in \mathbb{C}_i^{\perp}$ one has $[i,u,v] = -2i\,\pi_{\perp}(uv)$ and $|\pi_{\perp}(uv)|^2 = |u|^2|v|^2 - |h(u,v)|^2$, where $h(u,v) = -\pi_{\mathbb{C}_i}(uv)$ is the Hermitian form of $\mathbb{C}_i^{\perp} \cong \mathbb{C}^3$ and $\pi_{\perp}(uv)$ is orthogonal to $u, iu, v, iv$. Writing $p_D = vv^{\dagger}$ with $v_1 > 0$ (Appendix A.3, dyads), the three off-diagonal $\mathbb{C}_i^{\perp}$-parts and the matrix $P_W p_D$ are explicit in $v$, and Hermitian Gram determinants give $\|\mathrm{Def}\|^2 = 72 \det(P_W p_D)$.

Hence $\|F\|^2 = 16\lambda^2 \|\mathrm{Def}\|^2 = 16 \cdot 2^{14} \cdot 72\, N_W \Delta_W \det(P_W p_D) = 9 \cdot 2^{21} N_W \mathcal{A}$ by Proposition 5.6. $\mathcal{A}$ is symmetric, which gives the other orderings. If $\mathcal{A} \neq 0$, then $\Delta_W \neq 0$ by Proposition 5.6, so $N_W \neq 0$ because $\Delta_W = N_W (\mathrm{Im}\,\beta)^2$; with the case $\Delta_W = 0$ this gives the last assertion. □

The factor $9 \cdot 2^{21} N(p_A + p_B + p_C)$ depends only on three of the atoms and is positive when $\mathcal{A} > 0$. For M the theorem gives $\|F\|^2 = 2^{24}/3^9$ for every choice of the three atoms in the first three places; for S, $2^{17}/3^5$ or $2^{19}/3^6$ as $p_A, p_B$ are or are not both among them.

# G Formalization

The accompanying Lean 4 development (`github.com/macgregor-sudo/albert-sequential-measurements`) is built with Lean[197] version v4.36.0 on Mathlib[198] v4.36.0, unmodified, in the namespace `AlbertSequential`; declarations are named here and in the margin notes without that prefix, except the names beginning with `LemmaB`, which lie in a top-level namespace of that name, and those beginning with `AlbertChallenge`, the compared theorems below (restated in a file that imports only Mathlib, and checked against the proofs with the Comparator tool) and their translation layer. Namespace names such as `LemmaA`, `LemmaB`, `BlockB` and `AppF` are internal labels of the development, not references to numbered items of this paper. Every declaration named in this paper depends only on the axioms `propext`, `Classical.choice` and `Quot.sound`. The margin note beside a numbered statement names the declarations that prove it and

[196] Ibid. §11.

[197] L. de Moura and S. Ullrich, The Lean 4 theorem prover and programming language, in *Automated Deduction – CADE 28*, Lecture Notes in Comput. Sci. **12699**, Springer, Cham, 2021, 625–635. doi:10.1007/978-3-030-79876-5_37.

[198] The mathlib Community, The Lean mathematical library, in *Proc. 9th ACM SIGPLAN Int. Conf. on Certified Programs and Proofs (CPP 2020)*, ACM, 2020, 367–381. doi:10.1145/3372885.3373824. arXiv:1910.09336.

says whether the Lean statement is the printed one, is proved by a different route, is narrower, or has a different form. The development's fifteen files that define the octonions and the Albert algebra and prove their elementary laws are adapted from an earlier formalization registered in Palomar, a public registry of Lean formalizations, as entry PALOMAR-2026-10-16-000013[199], and three of the development's files on general Jordan algebras, for the fundamental formula of the quadratic representation, for power associativity and for the multiplication, quadratic and triple operators, are adapted from four files of the Physlib library[200]. Registered Lean formalizations of related mathematics, from which no code is used, include Hurwitz's theorem with a construction of the octonions[201], the Koecher–Alfsen–Shultz theorem, the spectral theorem, the trace form and the Peirce decomposition of Euclidean Jordan algebras[202], and the classifications of sequential products on $H_N(\mathbb{R})$ and $H_N(\mathbb{C})$ together with Albert's power-associativity theorem for commutative Jordan rings carrying a real-module structure[203].

The comparison class is widened in the proofs. An *algebraic model* consists of a unital associative real $*$-algebra, self-adjoint idempotents and a real-linear functional $\varphi$ with $\varphi(1) = 1$, $\varphi(X^*) = \varphi(X)$ and $\varphi(X^*X) \geq 0$. Every sequential projective representation over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$ induces one, so excluding algebraic models excludes all representations. Some statements of the development are `Prop`-valued interfaces, named `...Contract`; they are hypotheses or targets inside the development, not axioms, and each one used for the results below is proved there.

### Twelve statements in a self-contained file

Twelve theorems are stated in a separate file whose definitions (octonions, the Albert algebra, words, the three kinds of representation and algebraic models, the invariant $\mathcal{A}$, the configurations M and S, special subalgebras and density-operator states) use Mathlib only, and the development proves them with exactly these types. In that file an atom is an idempotent of trace one, and a subalgebra is special if a unital injective Jordan homomorphism maps it into the complex Hermitian $d \times d$ matrices for some $d \geq 1$. The theorems say the following. (1) If finitely many atoms generate $J$, no real, complex or quaternionic representation and no algebraic model reproduces their word probabilities from any $\rho \in J$; this is the no-go assertion of Theorem 4.1. (2) For such atoms there are a length $L$ and a linear inequality on words of length at most $L$ that holds in every real, complex

[199] MacGregor, "Finite observations and recognition of the compact Albert algebra", note 143.

[200] T. O. Diem, in Physlib (`leanprover-community/physlib`), Lean library, commit `08a6c1b5`, files `PhyslibAlpha/ProbabilisticTheory/JordanOrderUnit/Quadratic/Fundamental.lean`, `../Power/Associative.lean`, `.../Operator.lean` and `.../Quadratic/Triple.lean`, copyright 2026 Tom Ole Diem, Apache License 2.0. `https://github.com/leanprover-community/physlib`.

[201] B. Ehrlich, `composition-algebras`: Euclidean composition algebras over $\mathbb{R}$ in Lean 4, repository `ehrlich-b/composition-algebras`, 2026; Palomar registry entries PALOMAR-2026-08-29-000015 (commit `e37e22b0`) and PALOMAR-2026-08-29-000018 (commit `98d4437c`), Apache License 2.0.

[202] Ehrlich, *`euclidean-jordan`*, note 117.

[203] Ehrlich, *`sequential-products-lean`*, note 116.

and quaternionic representation and fails by a fixed positive amount from every state (Theorem 4.1, last sentence, and Corollary 4.16(a)). (3) For four generating atoms no representation and no algebraic model reproduces from $e/3$ every word of length at most seven that begins with "yes" on the first test, nor from $p_A$ every word of length at most six (Theorem 6.1(a)). (4) For four generating atoms some representation on $\mathbb{C}^{12}$ reproduces every word of length at most six from $e/3$ (Proposition 6.2). (5) Hence seven is the least such length, $L_{\min} = 7$, where the representations range over Hilbert spaces in the lowest universe of Lean's type theory. (6) The invariant $\mathcal{A}$ is nonnegative, and positive exactly when four atoms generate $J$ (Theorem 5.5(i), (iii)). (7) $\mathcal{A} \le \frac{4}{19683}$ for all atoms, with equality at $\mathsf{M}$ (Proposition 5.12). (8) The rational configuration $\mathsf{S}$ consists of atoms and has $\mathcal{A} = \frac{1}{5832}$. (9) Three atoms never generate $J$, and from every state some complex representation on $\mathbb{C}^3$ reproduces their word probabilities (Proposition 5.1). (10) If four atoms have $\mathcal{A} = 0$, then from every state some complex representation on $\mathbb{C}^6$ reproduces their word probabilities. (11) For any finite family of atoms, each of the following is equivalent to their not generating $J$: some unital subalgebra containing them is special; there are $d \in \{6, 24\}$ and one family of orthogonal projections on $\mathbb{C}^d$ such that from every state a representation with these projections and a density-operator state reproduces every word; from at least one state some representation of one of the three kinds, or an algebraic model, on a space in the lowest universe reproduces every word (Theorem 4.1, (1)⇔(2)⇔(3)). Its second condition is stronger than the printed (2), whose dimension may depend on the state. With (1), which holds in every universe, and Theorem F.5 for speciality in the operator sense, the printed equivalence follows. (12) The atoms are exactly the primitive idempotents: the nonzero idempotents that are not a sum of two nonzero idempotents $q, r$ with $q \circ r = 0$.

## Other results proved in the development

The proofs of these twelve theorems also establish results not among them. The construction behind (9) places the three atoms in a copy of $H_3(\mathbb{C})$ and uses rank-one projections, and that behind (10) uses fixed rank-two projections (`fourPoint_zero_c6_representation`); the statements record only the Hilbert spaces. $\mathcal{A} = 0$ exactly when the atoms lie in a copy of $H_3(\mathbb{H})$, and after an automorphism $\mathcal{A}$ is a product of squared coordinates times the squared norm of an octonion associator (Theorem 5.5(ii), (iv), Appendix A.3). The maximiser is unique up to automorphism, and at a maximiser each atom gives every word in the other three tests its probability from $e/3$ (Proposition 5.12, Corollary 5.14). Every Jordan homomorphism from $J$ into an associative algebra is zero (Theorem F.5).

The development further proves the following, each with the declarations named in its margin note. For the criterion: the cyclic representation, null vectors and purity (Lemmas 4.3–4.5), lifts and the orbit span (Lemmas 4.6, 4.7), the reconstructed homomorphism (Proposition 4.8) and the explicit state (Lemma 4.9). For four points: Definition 5.2, Propositions 5.3, 5.6, 5.10, Corollaries 5.7, 5.11, Lemmas 5.8, 5.13 and the lemmas of Appendix A.3. For finite certificates: Theorems 6.1, 6.10, 6.15, Corollary 6.3, Lemmas 6.4, 6.5, 6.7, 6.8, 6.9, C.1, Propositions 6.2,

6.11, 6.16, 6.17, 6.19, 6.20 and C.2. For five tests: Theorem 7.2, Lemmas 7.1, 7.3, Proposition 7.4, Corollary 7.5 and the lemmas of Appendix E; the operational sentence of Corollary 7.5 is derived from (11) and the Helly number five. For instruments: Theorem 8.1, Corollaries 8.2, 8.7, Propositions 8.4, 8.6, 8.10 and Lemma 8.5. For the algebraic obstruction: the corner relations, Theorems F.3, F.5, F.6 and Corollary F.4. Also Lemma 3.1, Theorem 3.3(a), Proposition 4.14 in part, and Definitions 3.2, 3.4, 3.6.

### Where the Lean statement is narrower

Several of these are proved in a narrower form, as the margin notes say. Definition 3.2 is stated for finite-dimensional positive forms. Corollary 4.16 and Theorem 3.3 only in part (a); Proposition 4.14 only in part (a), whose inclusions are proved by explicit constructions; Corollary 7.5 in its operational sentence for spaces in the lowest universe, and the sentence $L_{\min} = 7$ of Theorem 6.1 and Proposition 6.2 for Hilbert spaces in the lowest universe. Lemma 3.1 is stated for $V = J$, the word transport of Lemma 4.9 for $J$ and its subalgebras, Proposition 5.3 for an explicit rank-three model and for $J$, and Corollary F.4 and Lemma F.2 for maps defined on all of $J$, and Lemmas 4.5 and 4.7 for tests and outcomes that are atoms of $J$. Theorem 5.9 is proved in its algebraic form only. Lemmas 6.7 and 6.8 are proved for the words of $\mathcal{W}_{28}$, and the formulas of Lemma C.1 are not stated as printed. For Proposition 6.11 Lean gives a model on $\mathbb{C}^{12}$ through length six, not the equal mixture of $\mathsf{Q}_{\pm}$. For Proposition 6.17, parts (b) and (c) are not formalized. Proposition 6.19 is proved in its first assertion, Proposition 6.20 in an abstract trace-norm setting, Lemma 8.5 as an existence statement and Proposition 8.6 without its last sentence in general. Smaller gaps in Lemmas 4.4, 5.8, A.1, A.5, A.7, A.9 and Proposition 4.8 are named in their notes.

### Where the Lean statement has a different form

For Theorems 6.15 and F.6 the Lean statement is for Lean's own list or polynomial, which is not proved equal to the printed one. The list $\mathcal{F}_A$ of Theorem 6.15 is defined in Lean from an encoding of the 1,673 relation applications summarised in Table 1; the formal proof checks its count and length histogram, and the separating functionals are supported on words of the lists and the empty word. In the formula of Theorem F.6, $F$ is defined in Lean as the sum of the nine placements of the polarisation of $\psi$; its identification with the coefficient $[s^2t^2u^2]$ is not a Lean theorem. In Theorem 4.1 the three conditions are stated in Lean in the forms given in its note; the printed equivalence follows from that statement together with the exclusion in every universe.

### Proofs by hand only, and proofs in Lean only

The following numbered results are proved by hand only: Theorem 3.3(b), Theorem 4.11 with Definition 4.10, Corollary 4.13, Proposition 4.14(b), (c), Corollary 4.16(b), Lemma B.2, Lemma 6.6, Proposition 6.13, Proposition 6.17(b), (c),

Proposition 6.21, and the parts named above of the results that are narrower in Lean. Theorem 6.15, whose printed proof is an outline, and the statement of Remark 6.18 are proved in Lean, as is the length-eleven bound for five tests stated in Section 7 (`FiveTest.five_short_e3`, `FiveTest.five_short_p0`). The second-order behaviour of equalities near a complex subalgebra (Section 6.5) is stated without proof here and proved in Lean (`AppF.cor_so_residual`); the law of a random atom seen in a complex subalgebra (Section 5.5) is stated without proof and not formalized. Proofs written for reading of these statements are in the companion paper, in the sections "Words of length eleven for every configuration", "No equality through five steps", "The reflected quadruple and its budget", "The second-order neighbourhood of a complex subalgebra" and "What one complex subalgebra sees", respectively.

### Where the routes differ

Below Lemma 6.5 the formal proof of Theorem 6.1 differs from the printed one. A finite package of relation instances at every ordering of $B, C, D$ fixes all 406 Gram entries (the *Gram table*) between the 28 path vectors as polynomials in the traces (`LengthSeven.LemmaA.lemmaA`); the identity of that table with the comparison form is checked entry by entry (`LengthSeven.TableModel.tableModel`); and Lemma 6.9 is checked as a ring identity modulo the trace relation. The walk reduction of Lemma 6.6 is not used. For M and S the squared norms of the vectors of Theorem 6.10 are evaluated from the same Gram table at the thirteen traces of the configurations, which are themselves computed in Lean (`BlockB.isTracesM_maximizerConfig`, `BlockB.isTracesS_fixedS`).

The algebraic tetrahedron statement keeps the corner hypotheses of Appendix F, including the required ordered instances of (R1)–(R3), and the conclusion $\Delta_j X_j = 0$ in a possibly nonunital associative target, with no positivity or nonzero transition-probability assumption. Reconstruction is stated with all histories, zero probabilities, arbitrary dimension and projection rank, and nonfaithful or non-normal states of the model. The existence of a finite Hermitian embedding is separate from Lemma 4.9, which transports a state through an embedding already supplied.