

Classification of bipartite observables I: stable and uniformly stable observables

Eduardo Dueñez and José Iovino

Department of Mathematics, The University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249, USA

When two quantum systems are prepared independently and each is driven to an idealized limit, the expectation of a bipartite observable may depend on which system is idealized first. We call an observable stable when it does not. We show that an observable is stable if and only if every idealized preparation of one system can be simulated to arbitrary accuracy, uniformly over all probes, by finite mixtures of actual preparations, provided that the probing system is idealized first, and that, on stable observables, two idealized preparations have a unique, symmetric independent product. We also show that, under a uniform form of stability, the number of preparations that n adaptive probes can tell apart at a fixed margin is bounded by a polynomial in n , and that responses can be learned with a bounded number of mistakes; otherwise, at some margin, protocols with n probes tell apart 2^n preparations. Finally, we show that the swap test, Hong–Ou–Mandel coincidences and discrete equality tests are uniformly stable, with bounds independent of the dimension, while order comparisons of discrete spectra unbounded above are unstable. The proofs rest on the classical fact that a single double limit condition appears, under different names, in topology (Grothendieck), functional analysis (Krivine–Maurey), algebra (Arens), probability (Simons) and combinatorics (Pták), and in model theory as Shelah’s stability; the independent product corresponds to Harrington’s lemma. Model theory supplies a common language for these conditions; we use it as a Rosetta stone, pairing each physical question with its mathematical formulation.

Mathematics Subject Classification (2020). Primary 81P15, 81P16, 47B07, 03C45; Secondary 03C20, 03C66, 03C98, 28A35, 46B08, 46B10, 46B50, 46L06, 46L30, 46N50, 68Q11, 68Q32, 81P10, 81P18, 81P45.

Keywords. Order of idealization, bipartite observables, tensor products of states, swap test, Hong–Ou–Mandel interference, simulation by mixtures, independent product, double limit property, weak compactness, Arens regularity, finitely additive measures, Harrington’s lemma, model-theoretic stability, decision trees, Littlestone dimension, ultrapowers.

Contents

1 Introduction: one property, many languages

3

Eduardo Dueñez: Eduardo.Duenhez@utsa.edu

José Iovino: jose.iovino@utsa.edu

2	Physics: two orders of idealization	9
2.1	Actual states and effective observables	10
2.2	Idealized states	10
2.3	The two orders of idealization	11
3	Analysis: stability is the double limit property	12
3.1	Historical remarks	13
4	Experiments: the swap test, interference and equality tests are stable	14
4.1	The minimal tensor product	14
4.2	Factorization through a Hilbert space	15
4.3	The swap operator, symmetrization, and Hong–Ou–Mandel interference	15
4.4	Equality and order of discrete spectra	16
4.5	The order projection	16
5	Topology: idealized responses are observables	17
5.1	Topology: weak compactness	17
5.2	Geometry: factorization through a reflexive space	19
5.3	Algebra: the two orders of idealization agree	19
5.4	Pure probes suffice	19
5.5	Historical remarks	20
6	Convexity: finite mixtures simulate idealized preparations	20
6.1	A dimension-free Monte Carlo bound	22
6.2	Historical remarks	23
7	Probability: Fubini’s theorem for finitely additive measures	23
7.1	Fubini’s theorem for means	23
7.2	Countably additive measures	24
7.3	Ultrafilter products	24
7.4	Means, mixtures and samples	24
7.5	Historical remarks	25
8	Independence: Harrington’s lemma and the independent product	25
8.1	Abstract independence	25
8.2	Harrington’s lemma as an exchange of limits	25
8.3	The independent product of idealized preparations	26
8.4	Historical remarks	27
9	Combinatorics: polynomial versus exponential growth	27
9.1	From ladders to decision trees	28
9.2	Ultrapowers and uniform stability	29
9.3	Counting the preparations resolved by adaptive probes	31
9.4	Affine approximation versus exponential growth	33
9.5	Learning: Littlestone’s mistake bound	34
9.6	Historical remarks	35
10	Order: stable observables are closed under norm limits and orthogonal sums	35
10.1	Topological and algebraic closure	35
10.2	The boundary of closure: meets and strong limits	36

1 Introduction: one property, many languages

Idealized limits, such as infinite energy, vanishing bandwidth or the thermodynamic limit, are routine, and when two are taken at once their order can matter, as in spontaneous symmetry breaking. We study the order of idealization for two quantum systems prepared independently and idealized separately, and determine when the expectation of a bipartite observable is independent of which system is idealized first. This condition controls three operational tasks: simulating idealized preparations by finite mixtures, combining them independently, and telling them apart and learning them with adaptive probes.

Let H_1 and H_2 be complex separable Hilbert spaces describing two quantum systems, which we call Alice's and Bob's. We write $\mathfrak{D}(H_i)$ for the set of density operators on H_i . When the two systems are prepared independently in states ρ and σ , the joint system is described by $\rho \otimes \sigma$, and a bounded observable C on $H_1 \otimes H_2$ has expectation

$$\varphi_C(\rho, \sigma) = \text{Tr}((\rho \otimes \sigma)C).$$

When the observable is a projection P , the number $\varphi_P(\rho, \sigma)$ is the probability of the outcome P . Only self-adjoint operators are measured, but we call every bounded C an observable; this costs nothing, since φ_{C^*} is the complex conjugate of φ_C .

Two orders of idealization. Asymptotic regimes are described by sequences of preparations (ρ_k) and (σ_l) : states of increasing energy, outputs of circuits of increasing depth, photons of increasingly narrow bandwidth. Such sequences need not converge to a density operator: the energy eigenstates of an oscillator, for instance, have no normalizable limit. Their limit points are idealized states, much as plane waves and position eigenstates are idealized states in textbook quantum mechanics (see Section 2, where we make this precise). Idealizing both systems requires two limits, which can be taken in either order. We call C *stable* if for all sequences (ρ_k) in $\mathfrak{D}(H_1)$ and (σ_l) in $\mathfrak{D}(H_2)$ we have

$$\lim_{l \rightarrow \infty} \lim_{k \rightarrow \infty} \varphi_C(\rho_k, \sigma_l) = \lim_{k \rightarrow \infty} \lim_{l \rightarrow \infty} \varphi_C(\rho_k, \sigma_l) \quad (1)$$

whenever both iterated limits exist. The term comes from model theory and Banach space theory, where the same condition is called stability [She71, KM81]; it is unrelated to dynamical stability or to stability of matter. Let us notice that the simplest instability is classical: let (e_k) and (f_l) be orthonormal sequences, for instance energy eigenstates of two identical oscillators, and let P project onto the span of the vectors $e_k \otimes f_l$ with $k < l$. The value $\varphi_P(e_k, f_l)$, with e_k short for $|e_k\rangle\langle e_k|$, is 1 above the diagonal and 0 on and below it, so the two iterated limits are 0 and 1: the event “Bob's system has the higher energy” has no well-defined probability once both energies are idealized to infinity.

Three operational questions. A preparation σ of Bob's system acts as a *probe* of Alice's system: it assigns to each preparation ρ the response $\varphi_C(\rho, \sigma)$. The order of idealization matters because idealized preparations are used in three ways:

- (i) *Simulation.* An idealized preparation is not available in the laboratory, but finite mixtures of actual preparations are. The question is whether a mixture can reproduce the response of the idealized preparation to arbitrary accuracy, with a single mixture serving for every probe, including idealized probes.

- (ii) *Independent combination.* Given idealized preparations of the two systems, the question is whether there is a well-defined expectation for the joint system that extends φ_C and is symmetric in the two systems.
- (iii) *Discrimination and learning.* Bob learns about an unknown preparation of Alice's system online, probe by probe: each probe is chosen according to the previous answers and answers whether the response lies above or below a threshold. A protocol with n probes has 2^n possible answer sequences, so in principle n probes can tell apart exponentially many preparations. The question is whether the number they tell apart grows exponentially or only polynomially in n , and whether Bob can learn the response of an unknown preparation online with a bounded number of mistakes.

In Theorems 6.1 and 8.2 we show that positive answers to (i) and (ii) are each equivalent to stability. For (i), idealized preparations can be simulated by finite mixtures, with the probing system idealized first, exactly when the observable is stable. For (ii), the two orders of idealization assign the same expectation to every pair of idealized preparations exactly when the observable is stable, and the common value is then the unique, symmetric independent combination. For (iii), in Theorem 9.9 we show that the growth is exponential in general, and that a uniform form of stability is exactly what makes it polynomial. In Theorem 9.6 we show that, for a self-adjoint observable, uniform stability, which we define in Section 9.2, is equivalent to the stability of the expectation on ultrapowers, where it becomes $\Phi_C^u(x, y) = \lim_{k, u} \varphi_C(x_k, y_k)$ for families of experiments $x = (x_k)$ and $y = (y_k)$ indexed by k . Fix a margin $\delta > 0$, and count only answers whose response lies at least $\delta/2$ away from the threshold. In Theorem 9.9 and Proposition 9.13 we show that if the observable C is uniformly stable, then protocols with n probes tell apart at most $\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{d}$ preparations, a polynomial in n whose degree d depends on δ , and that the response of an unknown preparation can be learned online with at most d mistakes, and that otherwise, at some margin, protocols with n probes tell apart 2^n preparations for every n , and no bound on the mistakes exists. Counted over infinite sets, polynomial growth is linear: in Rowbottom's original definition, a stable theory has at most λ types over λ parameters, the least possible, whereas a priori there can be 2^λ (see Remark 9.10). In Section 9 we show that binary decision trees connect uniform stability with these counts: uniform stability bounds their height at each margin by d , a counting argument of Sauer–Shelah type turns this bound into the polynomial, and the failure of uniform stability yields trees of every height, that is, protocols with n probes that realize all 2^n answers.

Experiments. In Section 4 we show that stability holds for several basic observables of quantum optics and quantum information:

- *Local and product measurements* $A \otimes B$ and $A \otimes I + I \otimes B$, operator-norm limits of finite sums of such products (the minimal tensor product), all compact (for instance, finite-rank) observables, such as Bell-state projections, and all observables when one of the systems is finite-dimensional (see Theorem 4.1 and Corollary 4.2).
- *The swap test*, whose acceptance probability $\frac{1}{2}(1 + \text{Tr } \rho\sigma)$ estimates the overlap of two unknown states [BCWdW01], together with the projections onto the symmetric and antisymmetric subspaces (see Section 4.3).
- *Hong–Ou–Mandel coincidences* [HOM87]: two photons with internal states ρ and σ meeting at a balanced beam splitter give a coincidence with probability $\frac{1}{2}(1 - \text{Tr } \rho\sigma)$ (see Section 4.3).

- *Discrete equality tests*, such as “equal photon numbers” or “equal energies” of two particles confined in identical boxes (see Proposition 4.6).

We also show that, for infinite-dimensional systems, the last three observables are not norm limits of finite sums of products (see Proposition 4.3), and that their stability has a simple reason, which we establish in Proposition 4.4: the expectation is an inner product, $\varphi_C(\rho, \sigma) = \langle u(\rho), v(\sigma) \rangle$ for bounded maps u and v into one Hilbert space (for the swap test, the overlap $\text{Tr } \rho\sigma$ is the Hilbert–Schmidt inner product of the two density operators). In Theorems 6.5 and 9.11 we show that this inner-product form also yields quantitative bounds independent of the dimension. With Hilbert space replaced by a reflexive Banach space, this form characterizes stability: an observable is stable exactly when its expectation is the pairing of a reflexive Banach space with its dual (see Theorem A, the operator form of a theorem of Iovino [Iov99b]).

By contrast, in Proposition 4.6 we show that comparisons such as “Bob’s system has more photons” or, as above, “Bob’s oscillator has the higher energy” are unstable whenever both quantities have discrete spectra unbounded above. This instability has nothing to do with noncommutativity: the order projection is diagonal in a product basis.

Summary of main results.

- (1) **Simulation by finite mixtures.** An observable is stable exactly when every idealized preparation of Alice’s system is reproduced to arbitrary accuracy, uniformly over all probes, by finite mixtures of actual preparations, with the probing system idealized first; with Alice’s system idealized first, simulation can succeed even for unstable observables (see Theorem 6.1 and Example 6.2). For the swap test, Hong–Ou–Mandel coincidences and equality tests, equal-weight mixtures of $\lceil 4/\varepsilon^2 \rceil$ preparations suffice, in every dimension (see Theorem 6.5).
- (2) **Randomization.** An observable is stable exactly when averages over random preparations of the two systems can be taken in either order (a Fubini theorem; see Theorem 7.1), including random choices among idealized preparations, which are described by finitely additive measures; ordinary (countably additive) randomization of either preparation never detects instability (see Proposition 7.2).
- (3) **Independent combination.** An observable is stable exactly when the two orders of idealization assign the same expectation to every pair of idealized preparations; the common value is then the unique independent product, which is a state, extends the product of density operators and is symmetric in the two systems (see Theorem 8.2).
- (4) **Discrimination and learning.** A self-adjoint observable is *uniformly stable*, that is, its ladders have bounded length at every margin, exactly when its decision trees have bounded height at every margin, and this is equivalent to three other conditions: at every margin, the number of preparations resolved by n probes grows only polynomially in n ; at every margin, online learning of the response has a finite mistake bound, uniform in the threshold; and the expectation remains stable when preparations and probes are replaced by families of experiments indexed by k , with the expectation taken as a limit over k along an ultrafilter. Otherwise, at some margin, the observable encodes decision trees of every height, so that n adaptive probes resolve 2^n preparations. Uniform stability is strictly stronger than stability, as a comparison of levels in spin multiplets shows (see Theorem 9.6).

and Example 9.8). For the swap test and Hong–Ou–Mandel coincidences, trees at margin δ have height at most $4/\delta^2$, and for equality tests at most $16/\delta^2$, by Novikoff’s perceptron bound (see Section 9).

- (5) **Algebraic and topological closure.** Stable observables are closed under linear combinations, adjoints and norm limits, and include all norm limits of sums of products and all inner-product observables. Stable yes/no tests are closed under negation and under combining mutually exclusive tests. The closure stops at the quantum-logical “and” (the projection onto the intersection of two ranges) and at strong limits (see Section 10).

One property, many languages. The double limit condition (1) has been found independently in several parts of mathematics: as Grothendieck’s criterion for weak compactness [Gro52], as Pták’s combinatorial lemma on convex means [Ptá63], as a Fubini property for finitely additive measures [Ptá63, Sim68], as Arens regularity of bilinear forms [Are51, Pym65, Ü87], as Krivine and Maurey’s stability of Banach spaces [KM81], as Shelah’s stability [She71], as Harrington’s lemma on the symmetry of independence [HH84], and, through Shelah’s trees, as finite Littlestone dimension in learning theory [Lit88, CF19]. Iovino observed that Shelah’s order property is the two-valued case of the analytic condition and developed the real-valued theory [Iov94, Iov99a, Iov99b], building on Krivine and Maurey.

Model theory as a common language. Model theory provides a common language for these conditions. In continuous model theory [BYBHU08, BYU10], where the expectation $\varphi_C(\rho, \sigma)$ is a formula, each of the notions above has a standard counterpart, which we describe in Section 8 and in the historical remarks at the end of the sections. Idealized preparations correspond to types over a model, realized in an ultrapower of the model, and effective observables to types realized in the model itself. Weak compactness of idealized responses corresponds to definability of types, the two orders of idealization to the two orders of realizing coheirs, means on state spaces to Keisler measures, the agreement of the two orders to Harrington’s lemma, ladders and decision trees to Shelah’s order property and binary trees, and families of experiments to ultraproducts. Stability of a formula is the property that the double limit condition expresses in each case. This is the sense in which model theory serves as a Rosetta stone: in Table 1, each row states one dichotomy as a physical question, next to its mathematical language, its model-theoretic reading and the section where it is proved, and the model-theoretic column is the one through which any row can be translated into any other. Theorems A and B below state the equivalences precisely.

Theorem A (stability in seven languages). For every bounded observable C on $H_1 \otimes H_2$, the following are equivalent:

- (A1) *Analysis: order of idealization* (Krivine–Maurey, Shelah). The observable C is stable: the two iterated limits (1) agree. Definition 3.1
- (A2) *Topology: idealized responses* (Grothendieck). The effective observables $\text{Tr}_1((\rho \otimes I)C)$, for $\rho \in \mathfrak{D}(H_1)$, form a relatively weakly compact set; equivalently, the expectation extends to a separately continuous function of two idealized states. Theorem 5.1

THE DICHOTOMY

Physical question <i>language; model theory</i> (section)	Tame side: stable	Wild side: unstable
Experiments (§4)	swap test, Hong–Ou–Mandel coincidences, equality tests, product and compact observables	energy comparison $[A < B]$ with unbounded spectra, order projection $P_{k < l}$
DICHOTOMY I: STABILITY (see Theorem A)		
Order of idealization <i>analysis; order property</i> (§3)	the two iterated limits commute	infinite ladders: the iterated limits are 1 and 0
Idealized responses <i>topology; definable types</i> (§5.1)	effective observables relatively weakly compact; separately continuous extension	an idealized response that is not an observable
Factorization <i>geometry; reflexive pairing</i> (§5.2)	expectation is the pairing of a reflexive space E with E^*	no factorization through a reflexive space
Idealized states <i>algebra; coheirs in two orders</i> (§5.3)	Arens regular: $\omega_1 \otimes_A \omega_2 = \omega_1 \otimes_B \omega_2$	the two orders of idealization disagree
Simulation <i>convexity; definitions by averages</i> (§6)	idealized preparations simulated by finite mixtures, with the probing system idealized first	simulation error up to 1 for the order projection
Randomization <i>probability; Keisler measures</i> (§7)	Fubini for all finitely additive means	Fubini fails for some means, never for countably additive ones
Independent combination <i>independence; Harrington’s lemma</i> (§8)	a unique, symmetric independent product $\omega_1 \otimes \omega_2$	the two orders of realization disagree
DICHOTOMY II: UNIFORM STABILITY (see Theorem B)		
Uniform order of idealization <i>analysis; order property at a margin</i> (§9.1)	ladders of bounded length at every margin	ladders of every length at some margin
Decision trees <i>combinatorics; Shelah’s trees</i> (§9.1)	trees of bounded height at every margin	trees of every height at some margin
Families of experiments <i>ultraproducts; stability in all models</i> (§9.2)	expectation on ultrapowers is stable	infinite decision trees in ultrapowers
Discrimination <i>growth; Rowbottom’s λ-stability</i> (§9.3)	at margin δ , n probes resolve $O(n^d)$ preparations, $d = d_C(\delta)$	at some margin, n probes resolve 2^n preparations
Learning <i>learnability; Littlestone’s mistake bound</i> (§9.5)	finite mistake bound in the threshold game; at most $4/\delta^2$ for the swap test (perceptron)	no finite mistake bound at some margin

Table 1: Model theory as a Rosetta stone. Each row states the dichotomy as a physical question, with its mathematical language, its model-theoretic reading and the section where it is proved; the rows follow the order of the sections. For a bounded observable, either all the tame properties of Dichotomy I hold or all fail (see Theorem A), and likewise for Dichotomy II for self-adjoint observables (see Theorem B). The tame side of Dichotomy II is strictly smaller than that of Dichotomy I (see Example 9.8). The entries in the column on the wild side give the failure in its sharpest example, not in general.

- (A3) *Geometry: factorization* (Iovino; Davis–Figiel–Johnson–Pełczyński). The expectation is the pairing of a reflexive Banach space with its dual. Theorem 5.1
- (A4) *Algebra: idealized states* (Arens). The two orders of idealization agree on all idealized states. Theorem 5.1
- (A5) *Convexity: simulation* (Mazur). Every idealized preparation of Alice’s system is simulated to arbitrary accuracy, uniformly over all probes, by finite mixtures of actual preparations, with the probing system idealized first. Theorem 6.1
- (A6) *Probability: randomization* (Simons, Pták). Iterated integrals of the expectation against any two finitely additive probability measures on the state spaces commute. Theorem 7.1
- (A7) *Independence: independent combination* (Harrington). For every pair of idealized preparations, the two orders of realizing the pair assign the same value to C , which is then the unique, symmetric independent product. Theorem 8.2

Several of the equivalences in Theorem A are close at the level of proofs; for instance, (A4) and (A7) express the same identity, through Arens products and through Harrington’s lemma (see Section 8.3). Their interest lies in their different meanings.

Theorem B (uniform stability in five languages). For every bounded self-adjoint observable C on $H_1 \otimes H_2$, the following are equivalent:

- (B1) *Analysis: uniform order of idealization*. The observable C is uniformly stable: at every margin, its ladders, finite staircases like that of the order projection, have bounded length. Definition 9.5
- (B2) *Combinatorics: decision trees* (Shelah). At every margin, the binary decision trees encoded by C have bounded height. Theorem 9.6
- (B3) *Ultraproducts: families of experiments*. The expectation of C on ultrapowers, that is, on families of experiments, is stable. Theorem 9.6
- (B4) *Growth: discrimination* (Sauer–Shelah). For every margin $\delta > 0$, the number of preparations that n adaptive probes can tell apart at margin δ is bounded by a polynomial in n , instead of being 2^n . Theorem 9.9
- (B5) *Learnability: learning* (Littlestone). For every margin, the response of an unknown preparation can be learned, in the threshold game of Section 9.5, with a bounded number of mistakes. Proposition 9.13

Uniform stability implies stability, but not conversely (see Theorem 9.6 and Example 9.8). If $\varphi_C(\rho, \sigma) = \langle u(\rho), v(\sigma) \rangle$ with $\|u\|, \|v\| \leq 1$ in a Hilbert space, then both the degree of the polynomial in (B4) and the mistake bound in (B5) are at most $16/\delta^2$, in every dimension (see Theorem 9.11).

Polynomial approximation is affine. In the study of deep computations [ADIW24, DIMW⁺26], tame behavior corresponds, roughly, to approximation by polynomials in finitely many realized values, and wild behavior to exponential growth. For bipartite observables the approximating polynomials can be taken of degree one, and the reason is

convexity: since expectations are linear in the state, an affine combination $\sum_j t_j \varphi_C(\rho_j, \sigma) = \varphi_C(\sum_j t_j \rho_j, \sigma)$ of actual responses is itself the response of a physical mixture, so the actual responses form a convex set. For a stable observable, idealized responses lie in the weak closure of this set, and Mazur’s theorem turns weak approximation by a convex set into uniform approximation by its elements. Thus idealized responses are uniform limits of *affine* combinations of actual responses (see Theorem 6.1 and Remark 6.4). (The polynomial bounding the growth in (B4) has degree $d_C(\delta)$, which need not be one: for the comparison GT_n of Section 11, it is at least $n - 1$ at margin 1.) In Theorems 9.11 and 6.5 we show that, for the swap test, Hong–Ou–Mandel coincidences and equality tests, two further bounds come from linear arguments: Novikoff’s perceptron bound limits the heights of decision trees, and Maurey’s sampling argument limits the size of the mixtures, both independently of dimension.

Classical results and new results. For bounded functions on products of sets, the equivalences behind Theorem A are classical, as are the passage from ladders to trees, the characterization of uniform stability through ultraproducts, the Sauer–Shelah bound for trees, Littlestone’s mistake bound, Novikoff’s perceptron bound and Maurey’s sampling argument; we cite them in the historical remarks at the end of the sections. What the paper adds is their formulation for bipartite observables through the two orders of idealization, the physical meaning of each, and the following results, which we have not found in the literature in this form: the criterion that pure probes suffice (see Proposition 5.2); the fact that countably additive randomization never detects instability (see Proposition 7.2); the form of simulation by mixtures and its failure in the other order (see Theorem 6.1 and Example 6.2); the Monte Carlo bound (see Theorem 6.5); the independent product of idealized preparations (see Theorem 8.2); the strict gap between stability and uniform stability for observables (see Example 9.8); the dimension-free tree bound for Hilbert-factorable observables (see Theorem 9.11); and the failure of meets (see Proposition 10.3). Several of these are short consequences of the classical results, and we present them as such.

Organization of the paper. Each section is titled after the area it draws on. In Section 2 (physics) we make idealized preparations precise and introduce the two orders of idealization, and in Section 3 (analysis) we develop the double limit property. In Section 4 we treat the experiments listed above. In Sections 5–8 we prove Theorem A: in Section 5 the entries on topology, geometry and algebra (idealized responses, factorization and idealized states), in Section 6 the entry on convexity (simulation), in Section 7 the entry on probability (randomization), and in Section 8 the entry on independence (independent combination). In Section 9 (combinatorics) we prove Theorem B. With both dichotomies established, in Section 10 (order) we determine the closure properties of the class of stable observables and locate their boundary. Sections 11 and 12 contain an outlook and open questions.

2 Physics: two orders of idealization

The goal of this section is to give a precise meaning to an idealized preparation and to show that two idealized preparations can be combined in two different orders. Everything else in the paper compares these two orders.

2.1 Actual states and effective observables

We follow the mathematicians' convention that the inner product $\langle \psi, \chi \rangle$ is linear in ψ , and write matrix elements in Dirac notation. The space $\mathcal{T}(H)$ of trace-class operators on H , with the trace norm $\|\tau\|_1 = \text{Tr} |\tau|$, has dual $B(H)$, each operator $A \in B(H)$ acting by $\tau \mapsto \text{Tr}(\tau A)$. For a unit vector ψ , the pure state $|\psi\rangle\langle\psi|$ is often written simply ψ .

Definition 2.1 (Bilinear form and slice maps). Let $C \in B(H_1 \otimes H_2)$ be given.

- (1) The *bilinear form* of C is the map $m_C(\tau, \tau') = \text{Tr}((\tau \otimes \tau')C)$ on $\mathcal{T}(H_1) \times \mathcal{T}(H_2)$.
- (2) The *slice maps* of C are the operators $T_C: \mathcal{T}(H_1) \rightarrow B(H_2)$ and $S_C: \mathcal{T}(H_2) \rightarrow B(H_1)$ determined by

$$\text{Tr}(\tau' T_C(\tau)) = m_C(\tau, \tau') = \text{Tr}(\tau S_C(\tau')).$$

The restriction of m_C to $\mathfrak{D}(H_1) \times \mathfrak{D}(H_2)$ is the expectation φ_C , and both slice maps have norm at most $\|C\|$. In physical terms, the operator $T_C(\rho) = \text{Tr}_1((\rho \otimes I)C)$ is a partial trace: once Alice has prepared ρ , the observable C acts on Bob's system as the *effective observable* $T_C(\rho)$. In orthonormal bases (e_i) of H_1 and (f_k) of H_2 , and for a unit vector $\psi = \sum_i \psi_i e_i$, we have

$$\langle f_k | T_C(\psi) | f_l \rangle = \langle \psi \otimes f_k | C | \psi \otimes f_l \rangle = \sum_{i,j} \overline{\psi_i} \psi_j \langle e_i \otimes f_k | C | e_j \otimes f_l \rangle.$$

For a product $C = A \otimes B$, we have $T_C(\rho) = \text{Tr}(\rho A) B$: every preparation of Alice's system leaves a multiple of B on Bob's side, and the slice map T_C has rank one.

2.2 Idealized states

A *state* of $B(H)$ is a positive linear functional ω with $\omega(I) = 1$. The state space $\mathcal{S}(B(H)) \subseteq B(H)^*$ carries the weak* topology, in which it is compact. Density operators give the *normal*, or actual, states $A \mapsto \text{Tr}(\rho A)$; we call the remaining states *idealized*. For instance, if (e_k) is an orthonormal sequence (number states of an oscillator, say) and $\mathcal{U}$ is a nonprincipal ultrafilter on $\mathbb{N}$, then the state

$$\omega(A) = \lim_{k, \mathcal{U}} \langle A e_k, e_k \rangle$$

vanishes on every compact operator. It assigns probability zero to every finite set of levels and describes an oscillator excited to infinite energy. More generally, for a sequence (ρ_k) in $\mathfrak{D}(H)$, the *ultralimit state* $\lim_{k, \mathcal{U}} \rho_k$ is the state $A \mapsto \lim_{k, \mathcal{U}} \text{Tr}(\rho_k A)$.

Lemma 2.2. *The normal states are weak* dense in $\mathcal{S}(B(H))$. Every $\omega \in B(H)^*$ is a combination $\omega = c_1 \omega_1 - c_2 \omega_2 + i c_3 \omega_3 - i c_4 \omega_4$ of four states ω_j with $0 \leq c_j \leq \|\omega\|$.*

Proof. For the first statement, suppose that a state ω is not in the weak* closure of the convex set $\mathfrak{D}(H)$. By the Hahn–Banach theorem, there is a self-adjoint $A \in B(H)$ with $\omega(A) > \sup_{\rho \in \mathfrak{D}(H)} \text{Tr}(\rho A) = \sup \text{spec } A$, which is impossible for a state. The second statement is the Jordan decomposition of the hermitian parts of ω (see [KR83, Section 4.3]). $\square$

2.3 The two orders of idealization

Suppose that Alice's preparation is idealized to a state ω_1 of $B(H_1)$ and Bob's to a state ω_2 of $B(H_2)$. There are two natural ways of assigning an expectation to C .

Alice idealized first. Each actual state σ of Bob's system leaves on Alice's side the effective observable $S_C(\sigma)$, to which Alice's idealized preparation assigns the expectation $\omega_1(S_C\sigma)$. This is a bounded linear functional of $\sigma \in \mathcal{T}(H_2)$, so there is a unique operator $Z_{\omega_1} = S_C^*(\omega_1) \in B(H_2)$ with

$$\mathrm{Tr}(\sigma Z_{\omega_1}) = \omega_1(S_C\sigma) \quad \text{for all } \sigma \in \mathcal{T}(H_2).$$

The operator Z_{ω_1} is the effective observable that Alice's idealized preparation leaves on Bob's side. Bob then evaluates his state on it.

Bob idealized first. Symmetrically, Bob's idealized preparation leaves on Alice's side the operator $W_{\omega_2} = T_C^*(\omega_2)$ in $B(H_1)$, which satisfies $\mathrm{Tr}(\rho W_{\omega_2}) = \omega_2(T_C\rho)$ for all $\rho \in \mathcal{T}(H_1)$, and Alice evaluates her state on it.

Definition 2.3 (The two products). For states ω_1 of $B(H_1)$ and ω_2 of $B(H_2)$, and for $C \in B(H_1 \otimes H_2)$, we define

$$(\omega_1 \otimes_A \omega_2)(C) = \omega_2(Z_{\omega_1}), \quad (\omega_1 \otimes_B \omega_2)(C) = \omega_1(W_{\omega_2}).$$

The subscript names the system that is idealized first.

The next lemma explains the names: each product is an iterated limit, taken in one of the two orders.

Lemma 2.4. *Let (ρ_i) and (σ_j) be nets of normal states with weak* limits ω_1 and ω_2 , and let $C \in B(H_1 \otimes H_2)$ be given. Then we have*

$$(\omega_1 \otimes_A \omega_2)(C) = \lim_j \lim_i \varphi_C(\rho_i, \sigma_j), \quad (\omega_1 \otimes_B \omega_2)(C) = \lim_i \lim_j \varphi_C(\rho_i, \sigma_j).$$

Moreover, with $T_C^{**}: B(H_1)^* \rightarrow B(H_2)^{**}$ the bitranspose of T_C , we have

$$(\omega_1 \otimes_B \omega_2)(C) = \langle T_C^{**}(\omega_1), \omega_2 \rangle, \quad (\omega_1 \otimes_A \omega_2)(C) = \langle S_C^{**}(\omega_2), \omega_1 \rangle.$$

In particular, the map $\omega_2 \mapsto (\omega_1 \otimes_A \omega_2)(C)$ is weak* continuous for every C , being evaluation at the operator $Z_{\omega_1} \in B(H_2)$.

Proof. For a normal state σ , we have

$$\mathrm{Tr}(\sigma Z_{\omega_1}) = \omega_1(S_C\sigma) = \lim_i \mathrm{Tr}(\rho_i S_C\sigma) = \lim_i \varphi_C(\rho_i, \sigma),$$

and then $\omega_2(Z_{\omega_1}) = \lim_j \mathrm{Tr}(\sigma_j Z_{\omega_1})$. The formula for $\otimes_B$ is symmetric. Finally, we have $\langle T_C^{**}\omega_1, \omega_2 \rangle = \omega_1(T_C^*\omega_2) = \omega_1(W_{\omega_2})$, and symmetrically for S_C^{**} . $\square$

The two products are the two *Arens extensions* of the bilinear form m_C to the biduals [Are51], and *Arens regularity* of m_C is the statement that they agree. In Corollary 2.5 below, we show that when one of the two states is actual, they always agree.

Corollary 2.5 (Instability requires two idealized factors). *If ω_1 or ω_2 is normal, then $(\omega_1 \otimes_A \omega_2)(C) = (\omega_1 \otimes_B \omega_2)(C)$ for every $C \in B(H_1 \otimes H_2)$.*

Proof. If ω_1 is given by a density operator ρ , then we have $Z_{\omega_1} = T_C(\rho)$ and $(\omega_1 \otimes_B \omega_2)(C) = \mathrm{Tr}(\rho T_C^*\omega_2) = \omega_2(T_C\rho)$, so both products equal $\omega_2(T_C\rho)$. If ω_2 is given by a density operator σ , then we have $T_C^*(\sigma) = S_C(\sigma)$, and both products equal $\omega_1(S_C\sigma)$. $\square$

3 Analysis: stability is the double limit property

The goal of this section is to develop the double limit property for arbitrary bounded functions on products: its ultrafilter form, the ladders that witness its failure, and its preservation under the operations used later.

Definition 3.1 (Double limit property). Let X and Y be sets. A bounded function $f: X \times Y \rightarrow \mathbb{C}$ is *stable* if for all sequences (x_k) in X and (y_l) in Y , we have

$$\lim_{l \rightarrow \infty} \lim_{k \rightarrow \infty} f(x_k, y_l) = \lim_{k \rightarrow \infty} \lim_{l \rightarrow \infty} f(x_k, y_l)$$

whenever both iterated limits exist. An observable $C \in B(H_1 \otimes H_2)$ is *stable* if its expectation $\varphi_C(\rho, \sigma) = \text{Tr}((\rho \otimes \sigma)C)$ is stable on $\mathfrak{D}(H_1) \times \mathfrak{D}(H_2)$.

Two elementary facts are used throughout. First, every bounded array, and every finite family of bounded arrays, has subsequences of rows and columns along which both iterated limits exist: pass to a subsequence of columns along which every row converges (a diagonal argument), then to a subsequence of rows along which every remaining column converges, and then to further subsequences along which the two outer limits exist. Second, passing to subsequences does not change iterated limits that already exist.

The basic tool is the following extraction lemma, which turns a gap between two iterated ultrafilter limits into a *ladder*: an array that is close to one limit on and above the diagonal and close to the other below it.

Lemma 3.2 (Ladder extraction). Let $(a_{kl})_{k,l \in \mathbb{N}}$ be a bounded complex array, let $\mathcal{U}$ and $\mathcal{V}$ be nonprincipal ultrafilters on $\mathbb{N}$, and let

$$L_1 = \lim_{k, \mathcal{U}} \lim_{l, \mathcal{V}} a_{kl}, \quad L_2 = \lim_{l, \mathcal{V}} \lim_{k, \mathcal{U}} a_{kl}$$

be the two iterated limits. For every $\varepsilon > 0$ there are indices $k_1 < k_2 < \dots$ and $l_1 < l_2 < \dots$ such that $|a_{k_i l_m} - L_1| \leq \varepsilon$ for $i \leq m$ and $|a_{k_i l_m} - L_2| \leq \varepsilon$ for $i > m$.

Proof. The indices are chosen alternately, each in a set that belongs to the relevant ultrafilter. Put $\alpha_k = \lim_{l, \mathcal{V}} a_{kl}$ and $\beta_l = \lim_{k, \mathcal{U}} a_{kl}$. Given $k_1, l_1, \dots, k_{m-1}, l_{m-1}$, choose $k_m > k_{m-1}$ in the set

$$\{k : |\alpha_k - L_1| \leq \varepsilon/2 \text{ and } |a_{kl_i} - \beta_{l_i}| \leq \varepsilon/2 \text{ for all } i < m\},$$

which belongs to $\mathcal{U}$, and then $l_m > l_{m-1}$ in the set

$$\{l : |\beta_l - L_2| \leq \varepsilon/2 \text{ and } |a_{k_i l} - \alpha_{k_i}| \leq \varepsilon/2 \text{ for all } i \leq m\},$$

which belongs to $\mathcal{V}$. For $i \leq m$, we have $|a_{k_i l_m} - L_1| \leq |a_{k_i l_m} - \alpha_{k_i}| + |\alpha_{k_i} - L_1| \leq \varepsilon$, and for $i > m$, we have $|a_{k_i l_m} - L_2| \leq |a_{k_i l_m} - \beta_{l_m}| + |\beta_{l_m} - L_2| \leq \varepsilon$. $\square$

Corollary 3.3. Let X and Y be sets, and let $f: X \times Y \rightarrow \mathbb{C}$ be a bounded function. Then the function f is stable if and only if, for all sequences (x_k) in X and (y_l) in Y and all nonprincipal ultrafilters $\mathcal{U}$ and $\mathcal{V}$ on $\mathbb{N}$, we have

$$\lim_{k, \mathcal{U}} \lim_{l, \mathcal{V}} f(x_k, y_l) = \lim_{l, \mathcal{V}} \lim_{k, \mathcal{U}} f(x_k, y_l).$$

Proof. If ordinary iterated limits exist, they equal the ultrafilter ones; this gives one direction. For the other, suppose that the ultrafilter iterated limits are $L_1 \neq L_2$. Apply Lemma 3.2 to $a_{kl} = f(x_k, y_l)$ with $2\varepsilon < |L_1 - L_2|$, and pass to further subsequences (k_{i_p}) and (l_{m_q}) along which both ordinary iterated limits exist. For fixed p , we have $i_p \leq m_q$ for all large q , so the limit over q is within ε of L_1 , and hence so is the iterated limit $\lim_p \lim_q$. Symmetrically, the iterated limit $\lim_q \lim_p$ is within ε of L_2 . The two iterated limits differ, so the function f is unstable. $\square$

The argument after the application of Lemma 3.2 uses only the ladder inequalities. The same argument therefore shows the following: if, for a bounded function f on $X \times Y$ and complex numbers L_1 and L_2 , sequences (x_k) in X and (y_l) in Y and indices $k_1 < k_2 < \dots$ and $l_1 < l_2 < \dots$ give values $a_{k_i l_m} = f(x_{k_i}, y_{l_m})$ with $|a_{k_i l_m} - L_1| \leq \varepsilon$ for $i \leq m$ and $|a_{k_i l_m} - L_2| \leq \varepsilon$ for $i > m$, for some ε with $2\varepsilon < |L_1 - L_2|$, then the function f is unstable.

Lemma 3.4 (Preservation of stability). *Let X and Y be sets.*

- (a) (Sums with one-variable coefficients.) *Suppose that $f_1, \dots, f_r$ are stable on $X \times Y$ and $c_1, \dots, c_r: X \rightarrow \mathbb{C}$ and $d_1, \dots, d_r: Y \rightarrow \mathbb{C}$ are bounded. Then the function $\sum_j c_j(x) d_j(y) f_j(x, y)$ is stable. In particular, linear combinations of stable functions and separated sums $\sum_j c_j(x) d_j(y)$ are stable.*
- (b) (Uniform limits.) *If f and f' are bounded functions on $X \times Y$, the function f' is stable and $\sup_{X \times Y} |f - f'| \leq \varepsilon$, then any two iterated limits of f differ by at most 2ε . In particular, uniform limits of stable functions are stable.*
- (c) (Composition.) *If f is stable on $X \times Y$ and $g: X' \rightarrow X$ and $h: Y' \rightarrow Y$ are maps, then the function $(x', y') \mapsto f(g(x'), h(y'))$ is stable on $X' \times Y'$.*
- (d) (Reduction to states.) *For $C \in B(H_1 \otimes H_2)$, the expectation φ_C is stable on $\mathfrak{D}(H_1) \times \mathfrak{D}(H_2)$ if and only if the bilinear form m_C is stable on the product of the unit balls of $\mathfrak{T}(H_1)$ and $\mathfrak{T}(H_2)$.*

Proof. (a) Put $F = \sum_j c_j d_j f_j$, and let (x_k) and (y_l) be sequences along which both iterated limits of F exist. Pass to subsequences along which every $c_j(x_k)$ and every $d_j(y_l)$ converges, say to c_j^* and d_j^* , and both iterated limits of every f_j exist. By stability, the two iterated limits of f_j agree, say with Λ_j , and both iterated limits of F along the subsequences equal $\sum_j c_j^* d_j^* \Lambda_j$.

(b) Let L_1 and L_2 be iterated limits of f . Pass to subsequences along which both iterated limits of f' exist. They are equal, say to Λ , and we have $|L_1 - \Lambda| \leq \varepsilon$ and $|L_2 - \Lambda| \leq \varepsilon$.

(c) Sequences in X' and Y' are mapped to sequences in X and Y .

(d) One direction is trivial. For the other, decompose each τ in the unit ball of $\mathfrak{T}(H)$ as $\tau = c_1 \rho_1 - c_2 \rho_2 + i c_3 \rho_3 - i c_4 \rho_4$, using the positive and negative parts of the hermitian parts of τ , with $\rho_j \in \mathfrak{D}(H)$ and $0 \leq c_j \leq 1$ (choosing a fixed state when a part vanishes). These choices define maps $\tau \mapsto \rho_j(\tau)$ and bounded coefficients $\tau \mapsto c_j(\tau)$. The value $m_C(\tau, \tau')$ is then a sum of sixteen terms of the form $\pm c_j(\tau) c_{j'}(\tau') \varphi_C(\rho_j(\tau), \rho_{j'}(\tau'))$, with factors ± 1 or $\pm i$, so parts (a) and (c) apply. $\square$

3.1 Historical remarks

The double limit condition for a function on a product goes back to Grothendieck [Gro52], as a criterion for weak compactness in spaces of continuous functions. Independently,

Shelah introduced the *order property* of a formula and defined stable theories as those in which no formula has it [She71]; a ladder is a finite witness of the order property. Krivine and Maurey introduced stable Banach spaces, those whose norm satisfies the double limit condition along bounded sequences [KM81]. Iovino recognized that Shelah’s order property is the two-valued case of the same condition, developed stability for real-valued structures on the basis of Krivine–Maurey [Iov94, Iov99a], and proved that stable functions are exactly the restrictions of the pairing of a reflexive Banach space with its dual [Iov99b]. Ben Yaacov and Usvyatsov developed local stability in continuous model theory [BYBH08, BYU10], and Ben Yaacov derived definability of types directly from Grothendieck’s criterion [BY14]. Lemma 3.2 follows an argument in [Iov99b] (see also [Iov, Chapter 1]).

4 Experiments: the swap test, interference and equality tests are stable

The goal of this section is to decide stability for the observables listed in the introduction. We show that in these examples, stability comes from one of two sources, approximation by finite sums of products or factorization through a Hilbert space (see Theorem 4.1 and Proposition 4.4), and instability from an encoding of the order of $\mathbb{N}$ (see Theorem 4.7).

4.1 The minimal tensor product

We write $B(H_1) \odot B(H_2)$ for the space of finite sums of products $A \otimes B$. Its norm closure is the minimal C^* -tensor product $B(H_1) \otimes_{\min} B(H_2)$, which consists of the bipartite observables that are uniform limits of finite combinations of local ones.

Theorem 4.1. *Every $C \in B(H_1) \otimes_{\min} B(H_2)$ is stable. More generally, for every $C \in B(H_1 \otimes H_2)$, any two iterated limits of φ_C differ by at most $2 \operatorname{dist}(C, B(H_1) \odot B(H_2))$.*

Proof. For $C' = \sum_j A_j \otimes B_j$, we have $\varphi_{C'}(\rho, \sigma) = \sum_j \operatorname{Tr}(\rho A_j) \operatorname{Tr}(\sigma B_j)$, which is stable by Lemma 3.4(a), and $|\varphi_C - \varphi_{C'}| \leq \|C - C'\|$. Apply Lemma 3.4(b). $\square$

Corollary 4.2.

- (a) *Every compact operator on $H_1 \otimes H_2$, in particular every finite-rank projection such as a Bell-state projection, is stable.*
- (b) *If H_1 or H_2 is finite-dimensional, then every $C \in B(H_1 \otimes H_2)$ is stable.*

Proof. (a) The rank-one operators $|\psi \otimes \chi\rangle\langle\psi' \otimes \chi'| = |\psi\rangle\langle\psi'| \otimes |\chi\rangle\langle\chi'|$ span a dense subspace of the compact operators. (b) If H_1 has dimension n , then $C = \sum_{i,j \leq n} |e_i\rangle\langle e_j| \otimes C_{ij}$ for suitable $C_{ij} \in B(H_2)$, so the operator C lies in $B(H_1) \odot B(H_2)$. $\square$

Membership in the minimal tensor product is detected by the slice map.

Proposition 4.3. *If $C \in B(H_1) \otimes_{\min} B(H_2)$, then the slice map T_C is compact.*

Proof. The slice map $T_{A \otimes B} = \operatorname{Tr}(\cdot A)B$ has rank one, and $\|T_C - T_{C'}\| \leq \|C - C'\|$; norm limits of finite-rank operators are compact. $\square$

4.2 Factorization through a Hilbert space

In Section 4.3 and Proposition 4.6 we show that several physically basic observables lie outside the minimal tensor product and are nevertheless stable, because their expectation is an inner product.

Proposition 4.4. *Let X and Y be sets, let $\mathcal{H}$ be a Hilbert space, and let $u: X \rightarrow \mathcal{H}$ and $v: Y \rightarrow \mathcal{H}$ be bounded maps. Then the function $f(x, y) = \langle u(x), v(y) \rangle$ is stable. In particular, if $\varphi_C(\rho, \sigma) = \langle u(\rho), v(\sigma) \rangle$ for bounded maps $u: \mathfrak{D}(H_1) \rightarrow \mathcal{H}$ and $v: \mathfrak{D}(H_2) \rightarrow \mathcal{H}$, then the observable C is stable.*

Proof. Bounded sets in a Hilbert space are relatively weakly sequentially compact. Given sequences (x_k) and (y_l) along which both iterated limits exist, pass to subsequences with $u(x_{k_i}) \rightarrow \xi$ and $v(y_{l_j}) \rightarrow \xi'$. For each i , we have $\lim_j \langle u(x_{k_i}), v(y_{l_j}) \rangle = \langle u(x_{k_i}), \xi' \rangle$, which tends to $\langle \xi, \xi' \rangle$; the other iterated limit is also $\langle \xi, \xi' \rangle$. $\square$

4.3 The swap operator, symmetrization, and Hong–Ou–Mandel interference

Let K be a Hilbert space. The *swap* $F \in B(K \otimes K)$ is the unitary operator with $F(\psi \otimes \chi) = \chi \otimes \psi$. Since $F = F^*$ and $F^2 = I$, the operators

$$P_+ = \frac{1}{2}(I + F), \quad P_- = \frac{1}{2}(I - F)$$

are the projections onto the symmetric and the antisymmetric subspaces of $K \otimes K$.

Lemma 4.5 (Swap trick). *For all $\tau_1, \tau_2 \in \mathcal{T}(K)$, we have $\text{Tr}((\tau_1 \otimes \tau_2)F) = \text{Tr}(\tau_1 \tau_2)$.*

Proof. For $\tau_1 = |\psi\rangle\langle\psi'|$ and $\tau_2 = |\chi\rangle\langle\chi'|$, one computes $(\tau_1 \otimes \tau_2)F = |\psi \otimes \chi\rangle\langle\chi' \otimes \psi'|$, whose trace is $\langle\psi, \chi'\rangle\langle\chi, \psi'\rangle = \text{Tr}(\tau_1 \tau_2)$. Extend by bilinearity and continuity. $\square$

By Lemma 4.5, the expectation of the swap is $\varphi_F(\rho, \sigma) = \text{Tr}(\rho\sigma)$, the inner product of ρ and σ in the Hilbert space $S_2(K)$ of Hilbert–Schmidt operators, and therefore

$$\varphi_{P_\pm}(\rho, \sigma) = \frac{1}{2}(1 \pm \text{Tr}(\rho\sigma)).$$

By Proposition 4.4, the operators F and $P_\pm$ are stable. On the other hand, the slice map $T_F(\rho) = \rho$ is the inclusion $\mathcal{T}(K) \rightarrow B(K)$, which is not compact when K is infinite-dimensional, since the images of orthogonal pure states are at distance one; likewise $T_{P_\pm} = \frac{1}{2}(\text{Tr}(\cdot)I \pm T_F)$. By Proposition 4.3, the operators F and $P_\pm$ do not belong to $B(K) \otimes_{\min} B(K)$.

The same quantity appears in three experiments:

- *The swap test* [BCWdW01]. An ancilla controls a swap of two registers prepared in ρ and σ . The test accepts with probability $\frac{1}{2}(1 + \text{Tr} \rho\sigma) = \varphi_{P_+}(\rho, \sigma)$, which estimates the overlap.
- *Symmetric and antisymmetric subspaces.* For two independently prepared systems with the same state space, the number $\varphi_{P_\pm}(\rho, \sigma)$ is the probability that a projective test for exchange symmetry (respectively, antisymmetry) succeeds. For genuinely identical particles a product state $\rho \otimes \sigma$ is not itself admissible; the statement concerns distinguishable carriers, or the internal degrees of freedom of identical ones, as in the next item.
- *Hong–Ou–Mandel interference* [HOM87]. Two single photons whose internal (spectral and polarization) states are ρ and σ enter the two ports of a balanced beam splitter. The probability of a coincidence is $\frac{1}{2}(1 - \text{Tr} \rho\sigma) = \varphi_{P_-}(\rho, \sigma)$, which vanishes for equal pure states (bunching).

4.4 Equality and order of discrete spectra

Let A and B be self-adjoint operators on H_1 and H_2 with pure point spectra, with spectral decompositions $A = \sum_{\lambda} \lambda \Pi_{\lambda}^A$ and $B = \sum_{\lambda} \lambda \Pi_{\lambda}^B$. The events “equal values” and “Bob’s value is larger” are described by the projections

$$[A = B] = \sum_{\lambda} \Pi_{\lambda}^A \otimes \Pi_{\lambda}^B, \quad [A < B] = \sum_{\lambda < \lambda'} \Pi_{\lambda}^A \otimes \Pi_{\lambda'}^B.$$

Proposition 4.6 (Equality is stable, order is not). *Let A and B be as above.*

- (1) *The projection $[A = B]$ is stable. If A and B have infinitely many common eigenvalues, then the projection $[A = B]$ does not belong to $B(H_1) \otimes_{\min} B(H_2)$.*
- (2) *If A or B has finitely many eigenvalues, then the projection $[A < B]$ is stable.*
- (3) *If the point spectra of A and B are both unbounded above, then the projection $[A < B]$ is unstable.*

Proof. (1) We have $\varphi_{[A=B]}(\rho, \sigma) = \sum_{\lambda} \text{Tr}(\rho \Pi_{\lambda}^A) \text{Tr}(\sigma \Pi_{\lambda}^B) = \langle u(\rho), v(\sigma) \rangle_{\ell^2}$, where $u(\rho) = (\text{Tr} \rho \Pi_{\lambda}^A)_{\lambda}$ and $v(\sigma) = (\text{Tr} \sigma \Pi_{\lambda}^B)_{\lambda}$ are vectors of ℓ^2 -norm at most one, so Proposition 4.4 applies. For each common eigenvalue λ , pick a unit vector ψ_{λ} in the range of Π_{λ}^A . Then $T_{[A=B]}(\psi_{\lambda}) = \Pi_{\lambda}^B$, and these are mutually orthogonal nonzero projections, so the slice map $T_{[A=B]}$ is not compact; apply Proposition 4.3.

(2) If A has finitely many eigenvalues, then $[A < B] = \sum_{\lambda} \Pi_{\lambda}^A \otimes (\sum_{\lambda' > \lambda} \Pi_{\lambda'}^B)$ is a finite sum of products, and Theorem 4.1 applies. If B has finitely many eigenvalues, the same argument applies to $[A < B] = \sum_{\lambda'} (\sum_{\lambda < \lambda'} \Pi_{\lambda}^A) \otimes \Pi_{\lambda'}^B$.

(3) Choose eigenvalues $\lambda_1 < \lambda'_1 < \lambda_2 < \lambda'_2 < \dots$, with λ_k an eigenvalue of A with unit eigenvector ψ_k and λ'_l an eigenvalue of B with unit eigenvector χ_l . Then $\varphi_{[A < B]}(\psi_k, \chi_l)$ is 1 if $k \leq l$ and 0 if $k > l$, so the two iterated limits are 0 and 1. $\square$

4.5 The order projection

Theorem 4.7 (Instability of the order projection). *Let (e_k) and (f_l) be orthonormal sequences in H_1 and H_2 , and let $P_{k < l}$ be the projection onto the closed span of $\{e_k \otimes f_l : k < l\}$. Then $\varphi_{P_{k < l}}(e_k, f_l)$ is 1 if $k < l$ and 0 otherwise, so*

$$\lim_l \lim_k \varphi_{P_{k < l}}(e_k, f_l) = 0 \neq 1 = \lim_k \lim_l \varphi_{P_{k < l}}(e_k, f_l).$$

Hence the projection $P_{k < l}$ is unstable, and its distance to $B(H_1) \otimes_{\min} B(H_2)$ is at least $\frac{1}{2}$.

Proof. The values form the staircase

$$\begin{pmatrix} & f_1 & f_2 & f_3 & \cdots \\ e_1 & 0 & 1 & 1 & \cdots \\ e_2 & 0 & 0 & 1 & \cdots \\ e_3 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Each column tends to 0 and each row to 1. The bound on the distance follows from Theorem 4.1. $\square$

The table summarizes the examples; uniform stability is introduced in Section 9.2 below, and the entries in its column are justified by Theorem 9.6, Proposition 9.7 and Example 9.8. Uniform stability corresponds to what is called superstability in Banach space theory, where a space is superstable when all its ultrapowers are stable [Ray83]; we prefer the name uniform stability because superstability has a different meaning in model theory [She71].

Observable	Stable	Uniformly stable	Mechanism
$A \otimes B, A \otimes I + I \otimes B$	yes	yes	minimal tensor product
compact, Bell projections	yes	yes	minimal tensor product
any C , one factor finite-dimensional	yes	yes	finite sum of products
swap F ; P_+ (swap test); P_- (HOM)	yes	yes	factors through $S_2(K)$
$[A = B]$, discrete spectra	yes	yes	factors through ℓ^2
$[A < B]$, one spectrum finite	yes	yes	finite sum of products
sector comparison C_{sec} (see Example 9.8 below)	yes	no	reflexive, not superreflexive
$[A < B]$, spectra unbounded	no	no	encodes the order of $\mathbb{N}$
order projection $P_{k < l}$	no	no	encodes the order of $\mathbb{N}$

5 Topology: idealized responses are observables

The goal of this section is to prove that an observable is stable if and only if the response of Bob's system to every idealized preparation of Alice's is again an observable. In topological terms, this says that the slice map T_C , which sends Alice's preparation to the effective observable on Bob's side, maps bounded sets to relatively weakly compact sets. We then derive the geometric and algebraic forms of the same condition, and show that pure probes suffice to test it.

5.1 Topology: weak compactness

The proofs in this section and in Section 6 rest on five classical theorems about a Banach space X , which we recall in plain terms. A bounded operator is *weakly compact* if it maps the unit ball to a relatively weakly compact set.

- *Eberlein–Šmulian theorem* [Ebe47, Š40]. A subset of X is relatively weakly compact if and only if every sequence in it has a weak cluster point in X .
- *Gantmacher's theorem* [Gan40]. An operator $T: X \rightarrow Y$ between Banach spaces is weakly compact if and only if its bitranspose maps X^{**} into Y .
- *Mazur's theorem* [Maz33]. The weak closure and the norm closure of a convex subset of X coincide.
- *Krein's theorem* [KŠ40]. The closed convex hull of a relatively weakly compact subset of X is weakly compact.

- *Davis–Figiel–Johnson–Pełczyński theorem* [DFJP74]. Every weakly compact operator factors through a reflexive Banach space.

Theorem 5.1 (Characterization of stability). *For every $C \in B(H_1 \otimes H_2)$, the following are equivalent:*

- (1) *The observable C is stable.*
- (2) *The bilinear form m_C is stable on the product of the unit balls of $\mathcal{T}(H_1)$ and $\mathcal{T}(H_2)$.*
- (3) (Topology.) *The slice map $T_C: \mathcal{T}(H_1) \rightarrow B(H_2)$ is weakly compact.*
- (4) (Geometry.) *There are a reflexive Banach space E and bounded operators $a: \mathcal{T}(H_1) \rightarrow E$ and $J: E \rightarrow B(H_2)$ with $T_C = J \circ a$.*
- (5) (Algebra.) *For all states ω_1 of $B(H_1)$ and ω_2 of $B(H_2)$, we have $(\omega_1 \otimes_{\mathbf{A}} \omega_2)(C) = (\omega_1 \otimes_{\mathbf{B}} \omega_2)(C)$. Equivalently, for every state ω_1 of $B(H_1)$, we have $T_C^{**}(\omega_1) \in B(H_2)$, and then $T_C^{**}(\omega_1) = Z_{\omega_1}$.*
- (5') *The equality in (5) holds for all ultralimit states $\omega_1 = \lim_{k, \mathcal{U}} \rho_k$ and $\omega_2 = \lim_{l, \mathcal{V}} \sigma_l$ of sequences of normal states.*
- (6) (Separate continuity.) *There is a function $\hat{\varphi}: \mathcal{S}(B(H_1)) \times \mathcal{S}(B(H_2)) \rightarrow \mathbb{C}$, weak* continuous in each variable separately, with $\hat{\varphi}(\rho, \sigma) = \varphi_C(\rho, \sigma)$ for all normal states ρ and σ .*

We prove Theorem 5.1 in three steps, one for each language: in this subsection we show that (1) and (2) are equivalent and imply (3); in Section 5.2 we show that (3) and (4) are equivalent; and in Section 5.3 we close the circle through (5), (5') and (6).

Proof of the first step. We show that (1) and (2) are equivalent and imply (3). The equivalence of (1) and (2) is Lemma 3.4(d).

For the implication from (1) to (3), the idea is that a weak cluster point of a sequence of effective observables must be the operator Z_{ω_1} , and that any failure of this is a ladder. Let (ρ_k) be a sequence in $\mathfrak{D}(H_1)$, let $\mathcal{U}$ be a nonprincipal ultrafilter, and let $\omega_1 = \lim_{k, \mathcal{U}} \rho_k$. We claim that $\lim_{k, \mathcal{U}} \omega_2(T_C \rho_k) = \omega_2(Z_{\omega_1})$ for every state ω_2 . Suppose instead that the left side is L_1 and the right side is $L_2 \neq L_1$, and let $\varepsilon = |L_1 - L_2|/5$. By Lemma 2.4, we have $\lim_{k, \mathcal{U}} \text{Tr}(\sigma T_C \rho_k) = \text{Tr}(\sigma Z_{\omega_1})$ for every normal state σ . Choose recursively indices $k_1 < k_2 < \dots$ and normal states $\sigma_1, \sigma_2, \dots$ as follows. The index k_m is chosen in the set

$$\{k : |\omega_2(T_C \rho_k) - L_1| \leq \varepsilon \text{ and } |\text{Tr}(\sigma_i T_C \rho_k) - \text{Tr}(\sigma_i Z_{\omega_1})| \leq \varepsilon \text{ for all } i < m\},$$

which belongs to $\mathcal{U}$. The state σ_m is a normal state with $|\text{Tr}(\sigma_m T_C \rho_{k_i}) - \omega_2(T_C \rho_{k_i})| \leq \varepsilon$ for all $i \leq m$ and $|\text{Tr}(\sigma_m Z_{\omega_1}) - \omega_2(Z_{\omega_1})| \leq \varepsilon$; such a state exists by Lemma 2.2. The value $\varphi_C(\rho_{k_i}, \sigma_m)$ is then within 2ε of L_1 for $i \leq m$ and within 2ε of L_2 for $i > m$. Since $4\varepsilon < |L_1 - L_2|$, the remark after Corollary 3.3 shows that C is unstable, a contradiction. By Lemma 2.2, the claim extends from states to all of $B(H_2)^*$, so the operator Z_{ω_1} is a weak cluster point of $(T_C \rho_k)$. By the Eberlein–Šmulian theorem, the set $T_C(\mathfrak{D}(H_1))$ is relatively weakly compact. By the decomposition in the proof of Lemma 3.4(d), the image of the unit ball under T_C lies in a sum of four sets of the form $\{cA : |c| \leq 1, A \in T_C(\mathfrak{D}(H_1))\}$, each of which is relatively weakly compact, so the slice map T_C is weakly compact. $\square$

5.2 Geometry: factorization through a reflexive space

Condition (4) says that the expectation is the pairing of a reflexive space with its dual: with $u(\rho) = a(\rho)$ and $v(\sigma) = J^*(\sigma)$, the diagram

$$\begin{array}{ccc} \mathfrak{D}(H_1) \times \mathfrak{D}(H_2) & \xrightarrow{(u,v)} & E \times E^* \\ & \searrow \varphi_C & \downarrow \langle \cdot, \cdot \rangle \\ & & \mathbb{C} \end{array} \quad (2)$$

commutes.

Proof of the second step. We show that (3) and (4) are equivalent. If T_C is weakly compact, it factors through a reflexive space by the Davis–Figiel–Johnson–Pełczyński theorem, and then we have $\varphi_C(\rho, \sigma) = \text{Tr}(\sigma J(a\rho)) = \langle a(\rho), J^*(\sigma) \rangle$, which is the commutativity of (2). Conversely, the image under a of the unit ball is bounded, hence relatively weakly compact in the reflexive space E , and J is weak-to-weak continuous. $\square$

5.3 Algebra: the two orders of idealization agree

The algebraic form of stability is Arens regularity: the two products of Definition 2.3 agree. Its proof closes the circle of implications in Theorem 5.1, and it also shows that the agreement of the two products is equivalent to the separate continuity in (6), which returns to topology.

Proof of the third step. We show that (3) and (5) are equivalent, that (5) implies (5'), which implies (1), and that (5) and (6) are equivalent; together with the first two steps, this proves the theorem. We first show that (3) and (5) are equivalent. For a state ω_1 and a normal state σ , we have $\langle T_C^{**}\omega_1, \sigma \rangle = \omega_1(T_C^*\sigma) = \omega_1(S_C\sigma) = \text{Tr}(\sigma Z_{\omega_1})$. Hence the functional $T_C^{**}(\omega_1)$ belongs to $B(H_2)$ if and only if $T_C^{**}(\omega_1) = Z_{\omega_1}$, and by Lemma 2.4, this happens if and only if $(\omega_1 \otimes_B \omega_2)(C) = \omega_2(Z_{\omega_1}) = (\omega_1 \otimes_A \omega_2)(C)$ for all states ω_2 . Since states span $B(H_1)^*$ (see Lemma 2.2), condition (5) says that $T_C^{**}(B(H_1)^*) \subseteq B(H_2)$, which is (3) by Gantmacher's theorem.

The implication from (5) to (5') is trivial. For the implication from (5') to (1), suppose that C is unstable. There are sequences whose ordinary iterated limits differ, and by Lemma 2.4, the ultralimit states along any nonprincipal $\mathcal{U}$ and $\mathcal{V}$ reproduce these limits as the two products.

Finally, we show that (5) and (6) are equivalent. By Lemma 2.4, the maps $\omega_2 \mapsto (\omega_1 \otimes_A \omega_2)(C) = \omega_2(Z_{\omega_1})$ and $\omega_1 \mapsto (\omega_1 \otimes_B \omega_2)(C) = \omega_1(W_{\omega_2})$ are weak* continuous, so under (5) their common value is separately continuous. Conversely, let $\hat{\varphi}$ be as in (6). For a normal state σ and a net (ρ_i) of normal states with limit ω_1 , continuity in the first variable gives $\hat{\varphi}(\omega_1, \sigma) = \lim_i \varphi_C(\rho_i, \sigma) = (\omega_1 \otimes_A \sigma)(C)$. Both $\hat{\varphi}(\omega_1, \cdot)$ and $(\omega_1 \otimes_A \cdot)(C)$ are weak* continuous and normal states are dense, so we have $\hat{\varphi} = \otimes_A$. Symmetrically, we have $\hat{\varphi} = \otimes_B$, which is (5). $\square$

5.4 Pure probes suffice

In an experiment, the probes that are easiest to prepare are pure. The next proposition shows that pure probes detect every instability.

Proposition 5.2 (Pure probes suffice). *Let $C \in B(H_1 \otimes H_2)$ be given. If the restriction of φ_C to pairs of pure states is stable, then the observable C is stable.*

Proof. The idea is to represent effective observables as continuous functions on the closure of the pure states of $B(H_2)$, where Grothendieck's criterion needs to test only a dense set of points. For density operators we have $\varphi_{C^*} = \overline{\varphi_C}$, so the hermitian parts of C inherit the hypothesis, and we may assume that C is self-adjoint; then the operator $T_C(\rho)$ is self-adjoint for $\rho \geq 0$. Let $\Omega \subseteq \mathcal{S}(B(H_2))$ be the weak* closure of the pure states, a compact space. For a self-adjoint operator A , the function $\hat{A}(\omega) = \omega(A)$ on Ω satisfies $\sup_{\Omega} |\hat{A}| = \|A\|$, so the map $A \mapsto \hat{A}$ is a linear isometry of the self-adjoint part of $B(H_2)$ onto a closed subspace of $C(\Omega)$. Since the pure states are dense in Ω , restriction to the set P_2 of pure states of Bob's system is a linear isometry of $C(\Omega)$ onto a closed subspace of the space $\ell^\infty(P_2)$ of bounded functions on P_2 . By Grothendieck's criterion for bounded subsets of ℓ^∞ of a set [Gro52] (see also [Ptá63]), the double limit property of the hypothesis shows that the restrictions of the functions $\omega \mapsto \omega(T_C\psi)$, for pure states ψ of Alice's system, form a relatively weakly compact subset of $\ell^\infty(P_2)$, and hence the functions themselves form a relatively weakly compact subset of $C(\Omega)$. By the Hahn–Banach theorem, relative weak compactness passes to the closed subspace, so the set of operators $T_C\psi$, for ψ pure, is relatively weakly compact. By Krein's theorem, its closed convex hull is weakly compact. This hull contains $T_C(\mathfrak{D}(H_1))$, because every density operator is a trace-norm convergent convex combination of pure states. Arguing as at the end of the first step of the proof of Theorem 5.1, the slice map T_C is weakly compact. $\square$

5.5 Historical remarks

For general bounded bilinear forms, the equivalence of (2) and (5) in Theorem 5.1 is Pym's iterated limit criterion for Arens regularity [Pym65], and the equivalence of (3) and (5) is due to Ülger [Ü87]; the Arens products themselves are from [Are51]. The equivalence of (1) and (3) is an instance of Grothendieck's double limit criterion [Gro52]; the proof given here is a self-contained version adapted to states. The factorization (4) is the operator form of Iovino's theorem that stable functions are exactly restrictions of the pairing of a reflexive space with its dual [Iov99b, Theorem 3.3], whose proof combines the construction of Krivine–Maurey and Raynaud [KM81, Ray83] with the Davis–Figiel–Johnson–Pełczyński factorization [DFJP74].

6 Convexity: finite mixtures simulate idealized preparations

Idealized preparations are not available in the laboratory, but finite mixtures of actual ones are: Alice draws j with probability t_j and prepares ρ_j . The goal of this section is to show that stability is exactly the possibility of reproducing the response of an idealized preparation by such mixtures, uniformly over all probes, idealized ones included, and to bound the size of the mixtures for observables that factor through a Hilbert space. In Example 6.2 we show that the order of idealization matters.

Theorem 6.1 (Affine simulation). *For every $C \in B(H_1 \otimes H_2)$, the following are equivalent:*

- (1) *The observable C is stable.*
- (2) (Bob idealized first.) *For every state ω_1 of $B(H_1)$ and every $\varepsilon > 0$, there is $\rho \in \mathfrak{D}(H_1)$ such that for all states ω_2 of $B(H_2)$, we have*

$$|(\omega_1 \otimes_{\mathbf{B}} \omega_2)(C) - \omega_2(T_C(\rho))| \leq \varepsilon.$$

- (3) (Tail mixtures.) *For every sequence (ρ_k) in $\mathfrak{D}(H_1)$, every nonprincipal ultrafilter $\mathcal{U}$, every n and every $\varepsilon > 0$, there is a finite convex combination $\rho = \sum_{k \geq n} t_k \rho_k$ such that for all states ω_2 of $B(H_2)$, we have*

$$|(\omega_1 \otimes_{\mathbf{A}} \omega_2)(C) - \omega_2(T_C(\rho))| \leq \varepsilon, \quad \text{where } \omega_1 = \lim_{k, \mathcal{U}} \rho_k.$$

If these conditions hold, then the two products agree, and the density operator ρ in (2) can be taken to be a finite mixture from the tail of any net converging to ω_1 .

Proof. We first show that (1) implies (2) and (3). The idea is that, under stability, the idealized response is an actual operator in the weak closure of a convex set of actual responses, and Mazur's theorem converts weak approximation into norm approximation. By Theorem 5.1, we have $T_C^{**}(\omega_1) = Z_{\omega_1} \in B(H_2)$, and both products equal $\omega_2(Z_{\omega_1})$. If $\omega_1 = \lim_{k, \mathcal{U}} \rho_k$, then the state ω_1 lies in the weak* closure of the convex hull of $\{\rho_k : k \geq n\}$. The bitranspose T_C^{**} is weak*-to-weak* continuous, and on $B(H_2)$ the weak* topology of $B(H_2)^{**}$ is the weak topology. Hence the operator Z_{ω_1} lies in the weak closure of the convex set $T_C(\text{conv}\{\rho_k : k \geq n\})$, hence in its norm closure by Mazur's theorem: there is ρ in that convex hull with $\|Z_{\omega_1} - T_C(\rho)\| \leq \varepsilon$, and then $|\omega_2(Z_{\omega_1} - T_C(\rho))| \leq \varepsilon$ for every state ω_2 . The same argument applies to nets.

Proof that (2) implies (1). By Lemma 2.4, the functional $G = T_C^{**}(\omega_1) - T_C(\rho)$ on $B(H_2)^*$ satisfies $|G(\omega_2)| \leq \varepsilon$ for all states ω_2 , and hence we have $\|G\| \leq 4\varepsilon$ by Lemma 2.2. Hence the functional $T_C^{**}(\omega_1)$ lies within 4ε of $B(H_2)$ for every ε , hence in $B(H_2)$, and this holds for every state ω_1 . By Theorem 5.1(5), the observable C is stable.

Proof that (3) implies (1). Let (ρ_k) and (σ_l) be sequences along which the ordinary iterated limits

$$L_1 = \lim_k \lim_l \varphi_C(\rho_k, \sigma_l), \quad L_2 = \lim_l \lim_k \varphi_C(\rho_k, \sigma_l)$$

exist. Let $\mathcal{U}$ be a nonprincipal ultrafilter, let $\omega_1 = \lim_{k, \mathcal{U}} \rho_k$ and $\omega_2 = \lim_{l, \mathcal{U}} \sigma_l$, and put $\alpha_k = \lim_l \varphi_C(\rho_k, \sigma_l) = \omega_2(T_C \rho_k)$. Then we have $(\omega_1 \otimes_{\mathbf{A}} \omega_2)(C) = L_2$. Given $\varepsilon > 0$, choose n with $|\alpha_k - L_1| \leq \varepsilon$ for all $k \geq n$, and choose $\rho = \sum_{k \geq n} t_k \rho_k$ as in (3). Then we have

$$|L_2 - L_1| \leq |L_2 - \omega_2(T_C \rho)| + \left| \sum_k t_k \alpha_k - L_1 \right| \leq 2\varepsilon.$$

□

In Example 6.2 below, we show that the order of idealization in (2) and the restriction to tails in (3) cannot be dropped: if Alice is idealized first, approximation by actual states can succeed for an unstable observable.

Example 6.2 (The order of idealization matters). Let $(e_k)_{k \geq 0}$ and $(f_l)_{l \geq 1}$ be orthonormal bases of H_1 and H_2 , let S_k be the projection onto the span of $f_1, \dots, f_k$, and

put

$$P = |e_0\rangle\langle e_0| \otimes I + \sum_{k \geq 1} |e_k\rangle\langle e_k| \otimes S_k.$$

The operator P is a projection, and for $k \geq 1$ we have $\varphi_P(e_k, f_l) = 1$ if and only if $l \leq k$, so the projection P is unstable. On the other hand, the operator $T_P(\rho)$ is diagonal, with l -th diagonal entry $\rho_{00} + \sum_{k \geq l} \rho_{kk}$. These entries form a nonincreasing sequence in $[0, 1]$ whose first term is 1, and every such sequence (c_l) arises, from $\rho_{kk} = c_k - c_{k+1}$ and $\rho_{00} = \lim_l c_l$. The set of these diagonal operators is weak* closed, so for every state ω_1 , the operator Z_{ω_1} , which lies in the weak* closure of $T_P(\mathfrak{D}(H_1))$, equals $T_P(\rho)$ for some actual ρ . Thus $(\omega_1 \otimes_A \omega_2)(P) = \omega_2(T_P \rho)$ exactly, for all states ω_2 . For $\omega_1 = \lim_{k, \mathcal{U}} e_k$, the simulating state is e_0 ; but for $\omega_2 = \lim_{l, \mathcal{U}} f_l$, we have $(\omega_1 \otimes_B \omega_2)(P) = 0 \neq 1$.

Example 6.3 (Maximal failure for the order projection). Let $P = P_{k < l}$ with (e_k) an orthonormal basis, and let $\omega_1 = \lim_{k, \mathcal{U}} e_k$. The effective observable $T_P(e_k) = \sum_{l > k} |f_l\rangle\langle f_l|$ decreases to 0, so we have $Z_{\omega_1} = 0$ and $(\omega_1 \otimes_A \sigma)(P) = 0$ for every normal probe σ . For every $\rho \in \mathfrak{D}(H_1)$, however, we have $T_P(\rho) = \sum_k \rho_{kk} \sum_{l > k} |f_l\rangle\langle f_l|$, and $\text{Tr}(f_N T_P \rho) = \sum_{k < N} \rho_{kk}$ tends to 1. Thus no actual preparation reproduces ω_1 within any error less than 1, even on normal probes.

Remark 6.4 (Why degree one). In real-valued model theory, definitions of stable types are uniform limits of polynomials, or of lattice expressions, in instances of the formula [BYU10]. Since φ_C is affine in the state, an affine combination of instances is again an instance: $\sum_k t_k \varphi_C(\rho_k, \cdot) = \varphi_C(\sum_k t_k \rho_k, \cdot)$. Thus in Theorem 6.1 the approximating “formula” is a physical preparation, and polynomials of degree one suffice.

6.1 A dimension-free Monte Carlo bound

Mazur’s theorem gives no information on the size of the mixture. In Theorem 6.5 below, we show that when C factors through a Hilbert space, a sampling argument does: the mixture can be taken to be an empirical average of independent draws, and the law of large numbers in the Hilbert space bounds the error.

Let $\mathcal{H}$ be a Hilbert space, let $a: \mathfrak{T}(H_1) \rightarrow \mathcal{H}$ be a bounded linear map and $b: \mathfrak{T}(H_2) \rightarrow \mathcal{H}$ a bounded conjugate-linear map, and suppose that $m_C(\tau, \tau') = \langle a(\tau), b(\tau') \rangle$. We write $R_a = \sup_{\rho \in \mathfrak{D}(H_1)} \|a(\rho)\|$ and $\|b\| = \sup_{\|\tau'\|_1 \leq 1} \|b(\tau')\|$.

Theorem 6.5 (Monte Carlo simulation). *Let C , a and b be as above, let (ρ_k) be a sequence in $\mathfrak{D}(H_1)$, let $\mathcal{U}$ be a nonprincipal ultrafilter, and let $\omega_1 = \lim_{k, \mathcal{U}} \rho_k$. Then the observable C is stable, and for every n , every $\varepsilon > 0$ and every integer $M \geq (2R_a\|b\|/\varepsilon)^2$, there are indices $k_1, \dots, k_M \geq n$, not necessarily distinct, such that the equal-weight mixture $\bar{\rho} = \frac{1}{M} \sum_{j=1}^M \rho_{k_j}$ satisfies*

$$|(\omega_1 \otimes_A \omega_2)(C) - \omega_2(T_C \bar{\rho})| \leq \varepsilon \quad \text{for all states } \omega_2 \text{ of } B(H_2).$$

Proof. For $\xi \in \mathcal{H}$, let $B_\xi \in B(H_2)$ be the operator with $\text{Tr}(\tau' B_\xi) = \langle \xi, b(\tau') \rangle$ for all $\tau' \in \mathfrak{T}(H_2)$. The map $\xi \mapsto B_\xi$ is linear, we have $\|B_\xi\| \leq \|b\| \|\xi\|$, and $T_C(\rho) = B_{a(\rho)}$. Let ξ^* be the weak limit of $a(\rho_k)$ along $\mathcal{U}$. For a normal state σ , we have $\text{Tr}(\sigma Z_{\omega_1}) = \lim_{k, \mathcal{U}} \langle a(\rho_k), b(\sigma) \rangle = \langle \xi^*, b(\sigma) \rangle$, so we have $Z_{\omega_1} = B_{\xi^*}$. By Mazur’s theorem, there is a finite convex combination $\xi = \sum_{k \geq n} t_k a(\rho_k)$ with $\|\xi^* - \xi\| \leq \varepsilon/(2\|b\|)$. Let $\kappa_1, \dots, \kappa_M$ be independent random indices with $\Pr[\kappa_j = k] = t_k$. Since the summands are independent

with expectation ξ , we have

$$\mathbb{E} \left\| \xi - \frac{1}{M} \sum_{j=1}^M a(\rho_{\kappa_j}) \right\|^2 = \frac{1}{M} \mathbb{E} \|a(\rho_{\kappa_1}) - \xi\|^2 \leq \frac{R_a^2}{M},$$

so some realization $k_1, \dots, k_M$ satisfies $\|\xi - \frac{1}{M} \sum_j a(\rho_{k_j})\| \leq R_a/\sqrt{M} \leq \varepsilon/(2\|b\|)$. Hence $\|Z_{\omega_1} - T_C \bar{\rho}\| = \|B_{\xi^* - a(\bar{\rho})}\| \leq \varepsilon$. Stability follows from Proposition 4.4. $\square$

Corollary 6.6. *In the setting of Theorem 6.5, for the swap operator F and for the equality tests $[A = B]$, equal-weight mixtures of $\lceil 4/\varepsilon^2 \rceil$ preparations ρ_k with $k \geq n$ simulate the state ω_1 within ε , uniformly over all probes. For the swap test P_+ and the Hong–Ou–Mandel projection P_- , mixtures of $\lceil 1/\varepsilon^2 \rceil$ preparations suffice. These bounds do not depend on the dimension of the systems.*

Proof. For F , take $\mathcal{H} = S_2(K)$, $a(\tau) = \tau$ and $b(\tau') = \tau'^*$, so that $\langle \tau, \tau'^* \rangle_{S_2} = \text{Tr}(\tau' \tau)$; since $\|\tau\|_2 \leq \|\tau\|_1$, we have $R_a \leq 1$ and $\|b\| \leq 1$. On states, we have $T_{P_\pm}(\rho) = \frac{1}{2}(I \pm T_F(\rho))$, and correspondingly $Z_{\omega_1}^{\pm} = \frac{1}{2}(I \pm Z_{\omega_1}^F)$, so the error for $P_\pm$ is half the error for F . For $[A = B]$, take $\mathcal{H} = \ell^2$, $a(\tau) = (\text{Tr } \tau \Pi_\lambda^A)_\lambda$ and $b(\tau') = (\overline{\text{Tr } \tau' \Pi_\lambda^B})_\lambda$; then $\|a(\tau)\|_2 \leq \sum_\lambda |\text{Tr } \tau \Pi_\lambda^A| \leq \|\tau\|_1$, and similarly for b . $\square$

If T_C factors through a Banach space of Rademacher type $r > 1$, the same argument, with the type- r form of Maurey’s lemma [Pis81], gives mixtures of size $M = O(\varepsilon^{-r/(r-1)})$. Reflexivity alone, which characterizes stability, gives no rate (see Question (Q5)).

6.2 Historical remarks

That definitions of stable types in real-valued model theory are uniform limits of polynomials or lattice expressions in instances follows from the Stone–Weierstrass theorem (see [BYU10]); Ben Yaacov observed that, by Mazur’s theorem, averages of instances already suffice [BY14]. Pták’s lemma [Ptá63] is the classical source of convex means in the study of weak compactness. The sampling argument in the proof of Theorem 6.5 is Maurey’s empirical method, published by Pisier [Pis81].

7 Probability: Fubini’s theorem for finitely additive measures

The goal of this section is to show that stability is a Fubini property: the iterated integrals of the expectation against two probability measures on the state spaces commute. The property concerns *finitely additive* probability measures only, since for countably additive ones it always holds (see Proposition 7.2). We also relate the instability of the order projection to the noncommutativity of ultrafilter products.

7.1 Fubini’s theorem for means

A *mean* on a set Γ is a finitely additive probability measure defined on all subsets of Γ ; equivalently, it is a positive linear functional μ on the bounded functions on Γ with $\mu(1) = 1$. We write $\int g d\mu$ for $\mu(g)$. The *barycenter* of a mean μ on $\mathfrak{D}(H)$ is the state $\omega_\mu(A) = \int \text{Tr}(\rho A) d\mu(\rho)$ of $B(H)$.

Theorem 7.1 (Fubini's theorem for means). *An observable $C \in B(H_1 \otimes H_2)$ is stable if and only if for all means μ on $\mathfrak{D}(H_1)$ and ν on $\mathfrak{D}(H_2)$, we have*

$$\int \left(\int \varphi_C(\rho, \sigma) d\mu(\rho) \right) d\nu(\sigma) = \int \left(\int \varphi_C(\rho, \sigma) d\nu(\sigma) \right) d\mu(\rho).$$

Proof. The idea is that, since the expectation is affine in each state, iterated integrals depend only on barycenters, and so they are the two products of Section 2. Since $\varphi_C(\rho, \sigma) = \text{Tr}(\rho S_C \sigma)$ is affine in ρ , we have $\int \varphi_C(\rho, \sigma) d\mu(\rho) = \omega_\mu(S_C \sigma) = \text{Tr}(\sigma Z_{\omega_\mu})$, and integrating against ν gives $(\omega_\mu \otimes_A \omega_\nu)(C)$. Symmetrically, the other iterated integral is $(\omega_\mu \otimes_B \omega_\nu)(C)$. Every state ω is a barycenter: by Lemma 2.2, there is a net (ρ_i) of normal states with limit ω , and if $\mathcal{W}$ is an ultrafilter on its index set containing all tails, then $\mu(g) = \lim_{i, \mathcal{W}} g(\rho_i)$ defines a mean with barycenter ω . Thus the identity holds for all means if and only if condition (5) of Theorem 5.1 holds, that is, if and only if C is stable. $\square$

7.2 Countably additive measures

The means in Theorem 7.1 cannot be replaced by countably additive probability measures, because, as we show in Proposition 7.2 below, these never detect instability.

Proposition 7.2 (Countably additive measures never detect instability). *Let μ be a countably additive Borel probability measure on $\mathfrak{D}(H_1)$, with the trace norm, and let ν be a mean on $\mathfrak{D}(H_2)$. Then for every $C \in B(H_1 \otimes H_2)$, the two iterated integrals of φ_C against μ and ν agree.*

Proof. The barycenter of μ is the normal state given by the Bochner integral $\int \rho d\mu(\rho) \in \mathfrak{D}(H_1)$. By the proof of Theorem 7.1, the two iterated integrals are the two products of the barycenters, and these agree by Corollary 2.5. $\square$

Instability is therefore a purely finitely additive phenomenon: it requires that both systems be idealized by genuinely finitely additive averages.

7.3 Ultrafilter products

For nonprincipal ultrafilters $\mathcal{U}$ and $\mathcal{V}$ on $\mathbb{N}$, let $\mathcal{U} \otimes \mathcal{V}$ be the ultrafilter on $\mathbb{N}^2$ consisting of the sets $Q \subseteq \mathbb{N}^2$ with $\{k : \{l : (k, l) \in Q\} \in \mathcal{V}\} \in \mathcal{U}$. Let $\omega_1 = \lim_{k, \mathcal{U}} \rho_k$ and $\omega_2 = \lim_{l, \mathcal{V}} \sigma_l$. By Lemma 2.4, the product $(\omega_1 \otimes_B \omega_2)(C)$ is the limit of $\varphi_C(\rho_k, \sigma_l)$ along $\mathcal{U} \otimes \mathcal{V}$, and the product $(\omega_1 \otimes_A \omega_2)(C)$ is the limit along the transpose of $\mathcal{V} \otimes \mathcal{U}$. The tensor square of a single ultrafilter $\mathcal{W}$ is not symmetric: the set $\{(k, l) : k < l\}$ belongs to $\mathcal{W} \otimes \mathcal{W}$ but not to the transpose of $\mathcal{W} \otimes \mathcal{W}$. On the pairs (e_k, f_l) , the expectation of the order projection of Theorem 4.7 is the indicator of this set, so for $\omega_1 = \lim_{k, \mathcal{W}} e_k$ and $\omega_2 = \lim_{l, \mathcal{W}} f_l$ the two products are 1 and 0. Ultrafilters are $\{0, 1\}$ -valued finitely additive measures, and Proposition 7.2 explains why no countably additive version of this example exists.

7.4 Means, mixtures and samples

Finitely supported means on $\mathfrak{D}(H_1)$ correspond to the finite mixtures a laboratory can implement. In this language, Theorem 6.1 says that for a stable observable, a mean μ on $\mathfrak{D}(H_1)$ in Theorem 7.1 can be replaced by a finitely supported mean, with an error at most ε uniformly in ν , because the iterated integrals depend on μ only through its barycenter (see the proof of Theorem 7.1). Theorem 6.5 replaces Mazur's existence statement by an empirical average of independent draws.

7.5 Historical remarks

Pták’s combinatorial lemma [Ptá63] derives weak compactness from the existence of convex means that interchange limits, and the Fubini form for finitely additive means is studied by Simons [Sim68] and Levy [Lev07]. In model theory, finitely additive probability measures on definable sets, the Keisler measures [Kei87], admit two products according to the order of integration; the commutation of these products is tied to stability and generic stability [HPP08, HP11]. The two products $\otimes_A$ and $\otimes_B$ are their operator-algebraic counterparts (see the proof of Theorem 7.1).

8 Independence: Harrington’s lemma and the independent product

The goal of this section is to prove that on stable observables, two idealized preparations combine independently in a unique and symmetric way, while on an unstable observable the two orders of idealization disagree for some pair of idealized preparations (see Theorem 8.2 below). The mathematical counterpart is the theory of independence. In a stable theory, independence over a model is symmetric, and the key to symmetry is Harrington’s lemma: the definitions of two types over a model commute [HH84]. We recall the origins of abstract independence briefly, explain that Harrington’s lemma is an exchange of iterated ultrafilter limits, and then prove the physical statement.

8.1 Abstract independence

The axiomatic study of independence is older than model theory: it begins with the exchange argument of Steinitz for transcendence bases and with van der Waerden’s abstract dependence relations, and leads to matroids and pregeometries (see the historical remarks at the end of this section). In model theory, Shelah’s forking [She90] carries independence beyond pregeometries: nonforking independence is not given by a closure operator, yet in a stable theory it is symmetric and has existence and uniqueness of independent extensions over models. Lascar and Poizat [LP79] described it through heirs, coheirs and definability of types, and Harnik and Harrington [HH84] characterized stable theories as those carrying an abstract independence relation with these properties.

8.2 Harrington’s lemma as an exchange of limits

We recall the two-valued setting of [Iov, Chapter 4]. Let $\mathcal{M}$ be a model of a stable complete first-order theory, let p and q be complete types over $\mathcal{M}$ in finite tuples of variables x and y , and let $\varphi(x, y)$ be a formula with parameters from $\mathcal{M}$. Stability gives definitions $d_p\varphi(y)$ and $d_q\varphi(x)$ over $\mathcal{M}$: for every tuple b from $\mathcal{M}$, the formula $\varphi(x, b)$ belongs to p if and only if $d_p\varphi(b)$ holds in $\mathcal{M}$, and symmetrically for q . In Section 8.3 below, a type over $\mathcal{M}$ will correspond to an idealized state, and realizing a coheir to taking the inner limit first.

Fact 8.1 (Harrington’s lemma [HH84]). *In the setting above, we have $d_p\varphi(y) \in q$ if and only if $d_q\varphi(x) \in p$.*

Harrington’s lemma is the key step in the proof that nonforking independence is symmetric in stable theories, and hence in the Harnik–Harrington characterization [HH84]. Let us notice that Harrington’s lemma is an exchange of limits, which is the point for us: every type over $\mathcal{M}$ is finitely satisfiable in $\mathcal{M}$, so the type p is the trace of an ultrafilter $\mathcal{U}$ on the tuples from $\mathcal{M}$ of the length of x , in the sense that a formula $\psi(x)$ with parameters from $\mathcal{M}$ belongs to p if and only if the set of its solutions in $\mathcal{M}$ belongs to $\mathcal{U}$. Let $\mathcal{V}$ be an

ultrafilter with trace q , and regard φ as a $\{0, 1\}$ -valued function on pairs of tuples from $\mathcal{M}$. Then the two sides of Harrington's lemma take the form

$$d_p \varphi(y) \in q \iff \lim_{b, \mathcal{V}} \lim_{a, \mathcal{U}} \varphi(a, b) = 1, \quad d_q \varphi(x) \in p \iff \lim_{a, \mathcal{U}} \lim_{b, \mathcal{V}} \varphi(a, b) = 1,$$

so Harrington's lemma is exactly the exchange of the two iterated ultrafilter limits of Corollary 3.3, read in the language of types. The computation takes place entirely in $\mathcal{M}$ and involves no heirs, saturation or realizations. The same computation explains the two orders of realization: if b realizes the coheir of $\mathcal{V}$ and a then realizes the coheir of $\mathcal{U}$ over $\mathcal{M}b$, then the formula $\varphi(a, b)$ holds if and only if $\lim_{a', \mathcal{U}} \lim_{b', \mathcal{V}} \varphi(a', b') = 1$, and realizing the two coheirs in the other order exchanges the limits. Thus the two orders of realizing coheirs give the same type exactly when the iterated limits commute, and the symmetry of independence is the commutation of ultrafilter limits. The proofs of these statements, and of the Harnik–Harrington characterization, are given in [Iov, Chapter 4]; the real-valued versions, for formulas with values in a compact interval, are proved in Iovino's thesis [Iov94] (see also [BYU10]).

8.3 The independent product of idealized preparations

The correspondence with Section 2 is as follows. A type over a model corresponds to an idealized preparation, and the trace of an ultrafilter to an ultralimit state. Realizing Bob's coheir first and Alice's over it computes the inner limit over Bob's states, which corresponds to the product $\otimes_{\mathbf{B}}$; the other order corresponds to $\otimes_{\mathbf{A}}$. Under this correspondence, the conclusion of Harrington's lemma becomes condition (5) of Theorem 5.1. The analogy is one of form: the theorem below is proved directly from Theorem 5.1, without model theory. To state the resulting symmetry, let $\mathcal{O}$ be the set of stable observables in $B(H_1 \otimes H_2)$, let $F_{12}: H_1 \otimes H_2 \rightarrow H_2 \otimes H_1$ be the unitary with $F_{12}(\psi \otimes \chi) = \chi \otimes \psi$, and for states ω_1 of $B(H_1)$ and ω_2 of $B(H_2)$, let $\omega_1 \otimes \omega_2$ be the common value of the two products on $\mathcal{O}$:

$$(\omega_1 \otimes \omega_2)(C) = (\omega_1 \otimes_{\mathbf{A}} \omega_2)(C) = (\omega_1 \otimes_{\mathbf{B}} \omega_2)(C) \quad (C \in \mathcal{O}).$$

In Proposition 10.1 below, we show that the set $\mathcal{O}$ is a norm-closed self-adjoint subspace of $B(H_1 \otimes H_2)$ containing I .

Theorem 8.2 (The independent product). *For all states ω_1 of $B(H_1)$ and ω_2 of $B(H_2)$, the following hold:*

- (a) (Existence.) *The functional $\omega_1 \otimes \omega_2$ is a state of $\mathcal{O}$: it is linear, positive on positive elements of $\mathcal{O}$, and takes the value 1 at I .*
- (b) (Extension.) *If ω_1 and ω_2 are given by density operators ρ and σ , then $(\omega_1 \otimes \omega_2)(C) = \text{Tr}((\rho \otimes \sigma)C)$ for all $C \in \mathcal{O}$.*
- (c) (Uniqueness.) *For every $C \in \mathcal{O}$, the assignment $(\omega_1, \omega_2) \mapsto (\omega_1 \otimes \omega_2)(C)$ is separately weak* continuous, and it is the only separately weak* continuous function on $\mathcal{S}(B(H_1)) \times \mathcal{S}(B(H_2))$ that satisfies (b).*
- (d) (Symmetry.) *For every $C \in \mathcal{O}$, the observable $F_{12}CF_{12}^*$ on $H_2 \otimes H_1$ is stable, and $(\omega_2 \otimes \omega_1)(F_{12}CF_{12}^*) = (\omega_1 \otimes \omega_2)(C)$.*

Moreover, an observable $C \in B(H_1 \otimes H_2)$ is stable if and only if the two products $\otimes_{\mathbf{A}}$ and $\otimes_{\mathbf{B}}$ agree at C for all pairs of states.

Proof. (a) The slice map S_C is linear in C , and so is $(\omega_1 \otimes_{\mathbf{A}} \omega_2)(C) = \omega_2(S_C^* \omega_1)$. If $C \geq 0$ and $\sigma \geq 0$, then we have $S_C(\sigma) \geq 0$, so $\text{Tr}(\sigma Z_{\omega_1}) = \omega_1(S_C \sigma) \geq 0$ for every positive σ ; hence $Z_{\omega_1} \geq 0$ and $(\omega_1 \otimes_{\mathbf{A}} \omega_2)(C) = \omega_2(Z_{\omega_1}) \geq 0$. For $C = I$, we have $S_I(\sigma) = \text{Tr}(\sigma)I$ and $Z_{\omega_1} = I$.

(b) This is Corollary 2.5.

(c) This is the proof that (5) and (6) are equivalent in Theorem 5.1.

(d) Write $C' = F_{12} C F_{12}^*$. For density operators we have $\varphi_{C'}(\sigma, \rho) = \varphi_C(\rho, \sigma)$, so the observable C' is stable by Lemma 3.4(c). Let (ρ_i) and (σ_j) be nets of normal states with limits ω_1 and ω_2 . By Lemma 2.4, the product $(\omega_2 \otimes_{\mathbf{B}} \omega_1)(C')$, in which the second factor ω_1 is idealized first, equals $\lim_j \lim_i \varphi_{C'}(\sigma_j, \rho_i) = \lim_j \lim_i \varphi_C(\rho_i, \sigma_j) = (\omega_1 \otimes_{\mathbf{A}} \omega_2)(C)$. Both sides are the respective independent products.

The final statement is the equivalence of (1) and (5) in Theorem 5.1. $\square$

In the language of independence relations, (a) is the existence of an independent combination, (c) is its uniqueness (stationarity), and (d) is its symmetry. Outside $\mathcal{O}$ these properties are lost together: by Theorem 5.1(6), an unstable observable admits no separately continuous extension of its expectation to idealized states, and for the order projection the two orders of realization give 1 and 0 (see Theorem 4.7).

8.4 Historical remarks

Steinitz [Ste10] introduced transcendence bases of field extensions and proved, by the exchange argument that also gives the dimension of a vector space, that any two of them have the same cardinality. Van der Waerden [vdW30] observed that linear and algebraic dependence obey the same few rules, and derived the existence of bases and the invariance of their cardinality from an abstract dependence relation whose central axiom is exchange. Whitney [Whi35] isolated these rules as *matroids*, with equivalent axiomatizations by independent sets, bases, circuits and rank; Nakasawa developed the same theory independently in the same years [NK09], and Birkhoff [Bir35] and Mac Lane [ML36] related it to lattices and to projective geometry. Crapo and Rota [CR70] developed matroids as *combinatorial geometries*; in their terminology a closure operator with the exchange property is a *pregeometry*, the term still used in model theory. Baldwin and Lachlan [BL71] showed that algebraic closure in a strongly minimal set is a pregeometry, so that every model has a dimension. Kim and Pillay [KP97] extended the Harnik–Harrington characterization to simple theories. The symbol $\perp$ is due to Makkai [Mak84].

9 Combinatorics: polynomial versus exponential growth

The goal of this section is to prove Theorem B, which answers the question of discrimination: how many preparations n adaptive probes can tell apart, and whether responses can be learned. We show that unstable observables encode binary decision trees of every finite height, that bounded tree height is a uniform form of stability characterized through ultrapowers, that it is equivalent to polynomial growth of the number of preparations resolved, and that it is equivalent to a finite mistake bound in a learning game with a fixed threshold and margin. Throughout the section C is self-adjoint, so that φ_C is real; for a general C , the results apply to its hermitian parts. For $n \geq 0$, we write 2^n for the set of 0–1 sequences of length n and $2^{<n} = \bigcup_{m < n} 2^m$; we write $s \sqsubseteq \eta$ if s is an initial segment of η , and $s \frown \iota$ for s followed by $\iota \in \{0, 1\}$.

9.1 From ladders to decision trees

Definition 9.1 (Ladders and trees). Let $\delta > 0$ be given.

- (a) A δ -ladder of length N for C consists of states $\rho_1, \dots, \rho_N \in \mathfrak{D}(H_1)$ and $\sigma_1, \dots, \sigma_N \in \mathfrak{D}(H_2)$, a real number γ and a sign $\theta \in \{\pm 1\}$ such that

$$\theta(\varphi_C(\rho_i, \sigma_j) - \gamma) \geq \delta/2 \text{ if } i \leq j, \quad \theta(\varphi_C(\rho_i, \sigma_j) - \gamma) \leq -\delta/2 \text{ if } i > j.$$

Infinite ladders are defined in the same way.

- (b) A δ -tree of height n for C consists of states $\rho_\eta \in \mathfrak{D}(H_1)$ for $\eta \in 2^n$ (the *leaves*), states $\sigma_s \in \mathfrak{D}(H_2)$ for $s \in 2^{<n}$ (the *nodes*), a real number γ and a sign θ such that, whenever $s^\frown \iota \subseteq \eta$, we have

$$\theta(\varphi_C(\rho_\eta, \sigma_s) - \gamma) \geq \delta/2 \text{ if } \iota = 1, \quad \theta(\varphi_C(\rho_\eta, \sigma_s) - \gamma) \leq -\delta/2 \text{ if } \iota = 0.$$

A tree is an adaptive measurement protocol. Starting at the root, Bob probes with σ_s ; the expectation of C falls on one side of the threshold γ or the other, and Bob moves to $s^\frown 1$ or $s^\frown 0$ accordingly. After n probes, he knows which of the 2^n preparations ρ_η Alice made. Two leaves that first separate at the node s differ there by at least δ .

Proposition 9.2 (Instability gives infinite ladders). *Suppose that sequences (ρ_k) in $\mathfrak{D}(H_1)$ and (σ_l) in $\mathfrak{D}(H_2)$ give iterated limits $L_1 \neq L_2$ of $\varphi_C(\rho_k, \sigma_l)$. Then for every $\delta < |L_1 - L_2|$, there is an infinite δ -ladder for C whose states are taken from these sequences.*

Proof. For any nonprincipal ultrafilters $\mathcal{U}$ and $\mathcal{V}$ on $\mathbb{N}$, the iterated limits along $\mathcal{U}$ and $\mathcal{V}$ equal the ordinary iterated limits L_1 and L_2 . Apply Lemma 3.2 with $\varepsilon = (|L_1 - L_2| - \delta)/2$, and take $\gamma = (L_1 + L_2)/2$ and θ the sign of $L_1 - L_2$. $\square$

The passage from ladders to trees rests on a single arithmetic observation, which we prove in Proposition 9.3 below: binary sequences, ordered by their binary expansions, interleave so that each leaf lies to the left of every node at which it turns right, and to the right of every node at which it turns left (see Figure 9.1).

Proposition 9.3 (Ladders give trees). *Every δ -ladder of length $2^{n+1} - 1$ for C yields a δ -tree of height n for C , with the same γ and θ and with leaves and nodes taken from the ladder.*

Proof. For s of length at most n , put

$$\text{pos}(s) = \sum_{k < |s|} (1 - s(k))2^{-(k+1)} + 2^{-(|s|+1)} \in (0, 1).$$

The binary expansion of $\text{pos}(s)$ is $1 - s(0), \dots, 1 - s(|s| - 1), 1$, so the map pos is injective, and its image has $2^{n+1} - 1$ elements. We claim that if $s^\frown \iota \subseteq \eta$ with $\eta \in 2^n$, then we have $\text{pos}(\eta) < \text{pos}(s)$ if and only if $\iota = 1$. Indeed, with $k = |s|$, we have

$$\text{pos}(\eta) - \text{pos}(s) = -\iota 2^{-(k+1)} + \varrho, \quad \varrho = \sum_{k < j < n} (1 - \eta(j))2^{-(j+1)} + 2^{-(n+1)},$$

and $0 < \varrho \leq (2^{-(k+1)} - 2^{-n}) + 2^{-(n+1)} < 2^{-(k+1)}$. The bound $0 < \varrho < 2^{-(k+1)}$ is the point of the construction: the digits after position k can move $\text{pos}(\eta)$ by less than the digit at position k , so the turn taken at s alone decides on which side of $\text{pos}(s)$ the leaf lies.

Let $\rho_1, \dots, \rho_N$ and $\sigma_1, \dots, \sigma_N$, with $N = 2^{n+1} - 1$, be the states of the ladder, and let π be the increasing bijection from the image of pos onto $\{1, \dots, N\}$. Put $\rho_\eta = \rho_{\pi(\text{pos}(\eta))}$ for $\eta \in 2^n$ and $\sigma_s = \sigma_{\pi(\text{pos}(s))}$ for $s \in 2^{<n}$. If $s \frown \iota \subseteq \eta$ and $\iota = 1$, then $\pi(\text{pos}(\eta)) < \pi(\text{pos}(s))$, and the ladder gives $\theta(\varphi_C(\rho_\eta, \sigma_s) - \gamma) \geq \delta/2$; if $\iota = 0$, then $\pi(\text{pos}(\eta)) > \pi(\text{pos}(s))$, and the ladder gives the opposite inequality. $\square$

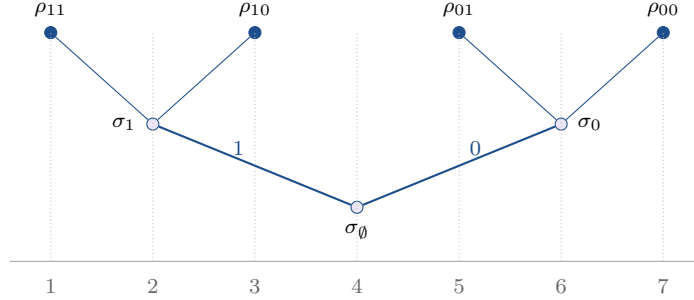


Figure 9.1: The map pos for $n = 2$. The seven sequences of length at most 2 are placed at the positions $1, \dots, 7$ of a ladder of length 7, ordered by pos ; the tree grows upward. Each leaf lies to the left of the nodes at which its branch turns right (label 1), so the ladder inequalities for $i \leq j$ are the tree inequalities for $\iota = 1$.

Theorem 9.4 (Unstable observables encode trees). *If C is unstable, then there is $\delta > 0$ such that for every n , there is a δ -tree of height n for C ; its leaves and nodes can be taken from any pair of sequences witnessing the instability, and in particular can be taken pure if those are pure.*

Proof. Combine Propositions 9.2 and 9.3. $\square$

For the order projection and for $[A < B]$ with unbounded spectra, the trees have $\delta = 1$ and consist of energy eigenstates (see Theorem 4.7, Proposition 4.6(3) and Proposition 9.3). The converse of Theorem 9.4 fails (see Example 9.8 below): trees of unbounded height detect a uniform form of instability, which we now describe through ultrapowers.

9.2 Ultrapowers and uniform stability

Fix a nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$. A sequence $x = (x_k)$ of states of Alice's system is an element of the ultrapower of her preparations; it should be thought of as a family of experiments indexed by k , in which both parties vary their preparations together. The *ultrapower expectation* of C is the function

$$\Phi_C^{\mathcal{U}}(x, y) = \lim_{k, \mathcal{U}} \varphi_C(x_k, y_k) \quad (x \in \mathfrak{D}(H_1)^{\mathbb{N}}, y \in \mathfrak{D}(H_2)^{\mathbb{N}}).$$

It is a joint limit, not an iterated one, and so it differs in kind from the idealized states of Section 2; it is the expectation of C on the Banach-space ultrapowers of $\mathcal{T}(H_1)$ and $\mathcal{T}(H_2)$ [Hei80]. Ladders and trees for $\Phi_C^{\mathcal{U}}$ are defined as in Definition 9.1, with elements of the ultrapowers in place of states. An *infinite δ -tree* consists of leaves x_η for every infinite branch $\eta \in 2^{\mathbb{N}}$ and nodes y_s for $s \in 2^{<\mathbb{N}}$ satisfying the inequalities of Definition 9.1(b) whenever $s \frown \iota \subseteq \eta$.

Definition 9.5. The observable C is *uniformly stable* if for every $\delta > 0$, the lengths of δ -ladders for C are bounded. A bounded observable C is *uniformly stable* if its hermitian parts $\frac{1}{2}(C + C^*)$ and $\frac{1}{2i}(C - C^*)$ are uniformly stable.

Theorem 9.6 (Uniform stability). *For every self-adjoint $C \in B(H_1 \otimes H_2)$, the following are equivalent:*

- (1) *The observable C is uniformly stable.*
- (2) *For every $\delta > 0$, the heights of δ -trees for C are bounded.*
- (3) *The ultrapower expectation $\Phi_C^{\mathcal{U}}$ is stable for every nonprincipal ultrafilter $\mathcal{U}$.*
- (4) *The ultrapower expectation $\Phi_C^{\mathcal{U}}$ is stable for some nonprincipal ultrafilter $\mathcal{U}$.*

If these conditions fail, then there is $\delta > 0$ such that for every nonprincipal $\mathcal{U}$, there is an infinite δ -tree for $\Phi_C^{\mathcal{U}}$. Every uniformly stable observable is stable.

Proof. The implication from (2) to (1) is Proposition 9.3.

Proof that (1) implies (3). Suppose that $\Phi_C^{\mathcal{U}}$ is unstable. The proof of Proposition 9.2 uses only Lemma 3.2, which holds for any bounded function on a product of sets, so it gives an infinite δ -ladder (x_i, y_j) for $\Phi_C^{\mathcal{U}}$. Fix N . For each of the N^2 pairs $i, j \leq N$, the set of k with $\theta(\varphi_C(x_{i,k}, y_{j,k}) - \gamma) \geq \delta/4$ (if $i \leq j$), respectively $\leq -\delta/4$ (if $i > j$), belongs to $\mathcal{U}$. For k in the intersection of these sets, the states $x_{i,k}$ and $y_{j,k}$ form a $\delta/2$ -ladder of length N for C . Hence the observable C is not uniformly stable.

The implication from (3) to (4) is trivial.

Proof that (4) implies (2). The idea is to glue trees of increasing height into an infinite tree in the ultrapower, and to observe that an infinite tree has $2^{\aleph_0}$ separated leaves, which no separable dual can accommodate. Let $\mathcal{U}$ be as in (4), and suppose that for some δ there are δ -trees $(\rho_\eta^{(n)}, \sigma_s^{(n)}, \gamma_n, \theta_n)$ of every height n . Pass to a set in $\mathcal{U}$ on which $\theta_n = \theta$ is constant, and let $\gamma = \lim_{n, \mathcal{U}} \gamma_n$. For an infinite branch $\eta \in 2^{\mathbb{N}}$, put $x_\eta = (\rho_{\eta|n}^{(n)})_n$, with $\eta|n$ the initial segment of length n , and for $s \in 2^{<\mathbb{N}}$, put $y_s = (\sigma_s^{(n)})_n$, with arbitrary entries for $n \leq |s|$. If $s \cap \iota \subseteq \eta$, then we have $s \cap \iota \subseteq \eta|n$ for all $n > |s|$, and passing to the limit, we have $\theta(\Phi_C^{\mathcal{U}}(x_\eta, y_s) - \gamma) \geq \delta/2$ if $\iota = 1$ and $\leq -\delta/2$ if $\iota = 0$. This is an infinite δ -tree for $\Phi_C^{\mathcal{U}}$, which proves the statement of Theorem 9.6 on infinite δ -trees.

Now suppose that $\Phi_C^{\mathcal{U}}$ is stable. After an affine rescaling to values in $[0, 1]$, which multiplies gaps by $c = 1/(2\|C\|)$, Iovino's theorem [Iov99b, Theorem 3.3] gives a reflexive Banach space E and maps $u: \mathfrak{D}(H_1)^{\mathbb{N}} \rightarrow E$ and $v: \mathfrak{D}(H_2)^{\mathbb{N}} \rightarrow E^*$, with values in the unit balls, such that $c(\Phi_C^{\mathcal{U}} + \|C\|) = \langle u(\cdot), v(\cdot) \rangle$. Let Y be the closed span in E^* of the countably many vectors $v(y_s)$. The space Y is separable and reflexive, so its dual Y^* is separable. If $\eta \neq \eta'$ first differ after the segment s , then $|\langle u(x_\eta) - u(x_{\eta'}), v(y_s) \rangle| \geq c\delta$, so the restrictions of the vectors $u(x_\eta)$ to Y are $2^{\aleph_0}$ points of Y^* at mutual distance at least $c\delta$, which is impossible in a separable space. Hence the function $\Phi_C^{\mathcal{U}}$ is unstable, contradicting (4).

Finally, an unstable observable has infinite ladders by Proposition 9.2, so it is not uniformly stable. $\square$

Theorem 9.6 applies verbatim to a bounded family of observables C_n on spaces $H_1^{(n)} \otimes H_2^{(n)}$, with $\Phi^{\mathcal{U}}(x, y) = \lim_{n, \mathcal{U}} \text{Tr}((x_n \otimes y_n)C_n)$; this is the relevant notion for registers of growing size (see Section 11).

Proposition 9.7. *Every self-adjoint $C \in B(H_1) \otimes_{\min} B(H_2)$ is uniformly stable, and so is every self-adjoint C with $\varphi_C(\rho, \sigma) = \langle u(\rho), v(\sigma) \rangle$ for bounded maps u and v into a Hilbert space. In particular, the swap operator, the projections $P_{\pm}$ and the equality tests $[A = B]$ are uniformly stable.*

Proof. For $C' = \sum_j A_j \otimes B_j$, the function $\Phi_{C'}^u(x, y) = \sum_j \lim_{k,u} \text{Tr}(x_k A_j) \cdot \lim_{k,u} \text{Tr}(y_k B_j)$ is a separated sum, and arguing as in the proof of Theorem 4.1, the function Φ_C^u is stable. In the Hilbert case, we have $\Phi_C^u(x, y) = \langle (u(x_k))_u, (v(y_k))_u \rangle$ in the ultrapower Hilbert space $\mathcal{H}^u$, and Proposition 4.4 applies. Conclude with Theorem 9.6. $\square$

Example 9.8 (Stable but not uniformly stable). Let both H_1 and H_2 be the Hilbert space $\bigoplus_{n \geq 1} \mathbb{C}^n$, with orthonormal bases $(e_i^n)_{i \leq n}$ of the summands; for instance, the summand $\mathbb{C}^{2j+1}$ is the spin- j representation, with basis $|j, m\rangle$ ordered by m . Put

$$C_{\text{sec}} = \sum_{n \geq 1} \sum_{1 \leq i \leq l \leq n} |e_i^n\rangle\langle e_i^n| \otimes |e_l^n\rangle\langle e_l^n|,$$

the projection onto “both systems lie in the same sector, and Alice’s level does not exceed Bob’s”. In the sector $\mathbb{C}^n$, the states $e_1^n, \dots, e_n^n$ form a 1-ladder of length n , so the projection C_{sec} is not uniformly stable and has 1-trees of every finite height. The projection C_{sec} is, however, stable: the effective observable $T_{C_{\text{sec}}}(\rho)$ depends only on the diagonal entries ρ_{ii}^n , and $T_{C_{\text{sec}}} = J \circ a$ with

$$a: \mathcal{T}(H_1) \rightarrow \left(\bigoplus_n \ell_n^1 \right)_{\ell^2}, \quad a(\rho) = ((\rho_{ii}^n)_{i \leq n})_n, \quad J((w_n)_n) = \bigoplus_n \sum_{l \leq n} \left(\sum_{i \leq l} w_{n,i} \right) |e_l^n\rangle\langle e_l^n|.$$

Here $\|a(\rho)\| \leq \sum_n \sum_i |\rho_{ii}^n| \leq \|\rho\|_1$ and $\|J\| \leq 1$, and the ℓ^2 -sum of finite-dimensional spaces is reflexive, so Theorem 5.1(4) applies.

The space $(\bigoplus_n \ell_n^1)_{\ell^2}$ is reflexive but not superreflexive. Superreflexivity is relevant here: if T_C factors through a superreflexive space E , then the ultrapower expectation factors through the reflexive space E^u [Hei80], so the observable C is uniformly stable. Whether the converse holds is Question (Q2).

9.3 Counting the preparations resolved by adaptive probes

A tree of height n lets n adaptive probes tell apart 2^n preparations. We now show that this exponential behavior is the only alternative to polynomial behavior.

Fix $\delta > 0$, a threshold γ and a sign θ . A *probe tree of depth n* is a family $(\sigma_s)_{s \in 2^{<n}}$ of states of Bob’s system. A branch $\eta \in 2^n$ is *realized* by $\rho \in \mathfrak{D}(H_1)$ if, whenever $s \frown \iota \subseteq \eta$, the inequalities of Definition 9.1(b) hold for $\varphi_C(\rho, \sigma_s)$. Realized branches are the outcomes that the adaptive protocol can return with margin δ , and two preparations that realize different branches are told apart by n probes. For a set $\mathcal{R} \subseteq \mathfrak{D}(H_1)$ of preparations, let $d(\mathcal{R}) \in \{-1, 0, 1, \dots, \infty\}$ be the largest height of a δ -tree with these γ and θ whose leaves lie in $\mathcal{R}$, with $d(\emptyset) = -1$. The *growth function* of C is

$$\text{gr}_C(n, \delta) = \sup \# \{ \eta \in 2^n : \eta \text{ is realized by some } \rho \in \mathfrak{D}(H_1) \},$$

the supremum being over γ, θ and probe trees of depth n , and the *tree dimension* $d_C(\delta) = \sup_{\gamma, \theta} d(\mathfrak{D}(H_1))$ is the largest height of a δ -tree for C . We write $\binom{n}{\leq d} = \sum_{i=0}^d \binom{n}{i}$.

Theorem 9.9 (Polynomial versus exponential growth). *Let $C \in B(H_1 \otimes H_2)$ be self-adjoint, and let $\delta > 0$ be given.*

(1) *If $d = d_C(\delta)$ is finite, then for all n we have*

$$\text{gr}_C(n, \delta) \leq \binom{n}{\leq d} \leq \left(\frac{en}{d}\right)^d \quad (n \geq d \geq 1),$$

which is a polynomial in n of degree d .

(2) *If $d_C(\delta)$ is infinite, then we have $\text{gr}_C(n, \delta) = 2^n$ for every n .*

Consequently, the observable C is uniformly stable if and only if $\text{gr}_C(\cdot, \delta)$ grows polynomially for every $\delta > 0$; otherwise it equals 2^n for some δ . In particular, unstable observables have exponential growth.

Proof. Statement (2) holds because a δ -tree of height at least n , restricted to its nodes of length less than n and one leaf above each branch of length n , realizes every branch of a probe tree of depth n . For (1), the idea is that at each probe, at most one of the two halves of the remaining preparations can carry a tree as tall as the whole set. Fix γ, θ and a probe tree of depth n . We show by induction on n that the number of branches realized by elements of a set $\mathcal{R} \subseteq \mathfrak{D}(H_1)$ is at most $\binom{n}{\leq d(\mathcal{R})}$, with the convention that $\binom{n}{\leq -1} = 0$. For $n = 0$ this is clear. For $n \geq 1$, let $\mathcal{R}_\iota$ be the set of elements of $\mathcal{R}$ that satisfy the inequality for ι at the root. The sets $\mathcal{R}_0$ and $\mathcal{R}_1$ are disjoint, because the two inequalities are separated by δ . A branch $\iota \cap \eta'$ is realized in $\mathcal{R}$ exactly when η' is realized in $\mathcal{R}_\iota$ for the subtree $(\sigma_\iota \cap s)_s$ of depth $n - 1$. We have $d(\mathcal{R}_\iota) \leq d(\mathcal{R})$, and the two numbers $d(\mathcal{R}_0)$ and $d(\mathcal{R}_1)$ cannot both equal $d(\mathcal{R})$: two δ -trees of height $d(\mathcal{R})$ with leaves in $\mathcal{R}_0$ and in $\mathcal{R}_1$, placed above the root probe σ_θ , would form a δ -tree of height $d(\mathcal{R}) + 1$ with leaves in $\mathcal{R}$. By induction and Pascal's rule, the number of realized branches is at most $\binom{n-1}{\leq d} + \binom{n-1}{\leq d-1} = \binom{n}{\leq d}$, with $d = d(\mathcal{R})$. The bound $\binom{n}{\leq d} \leq (en/d)^d$ is standard. The last statements follow from Theorems 9.4 and 9.6. $\square$

Remark 9.10 (Rowbottom's definition of stability). Stability has many equivalent characterizations, but its original definition was a growth condition. Rowbottom [Row64] called a complete theory T *stable in* an infinite cardinal λ if for every model $\mathcal{M}$ of T and every set $A \subseteq \mathcal{M}$ of cardinality λ , the space $S_1(A)$ of complete 1-types over A has cardinality at most λ . The reverse inequality is trivial, since distinct elements of A realize distinct types over A ; so stability in λ says that $|S_1(A)| = |A|$, that is, the number of types grows linearly with the number of parameters, whereas a priori it can be as large as 2^λ . For a single formula φ , stability is equivalent to the bound $|S_\varphi(A)| \leq |A|$ for every infinite set A of parameters (see [Pil96, Chapter 1] and [She90]). Theorem 9.9 is a finite and adaptive counterpart of this condition: the number of preparations resolved by n probes is bounded by a polynomial in n . For an infinite cardinal λ and an integer $d \geq 1$, we have $\lambda^d = \lambda$, so on infinite parameter sets the distinction between linear and polynomial bounds disappears.

In Theorem 9.11 below, we show that for observables that factor through a Hilbert space, the degree of the polynomial is controlled explicitly, by a linear learner.

Theorem 9.11 (Linear bound for Hilbert-factorable observables). *Let $C \in B(H_1 \otimes H_2)$ be self-adjoint, and suppose that $\varphi_C(\rho, \sigma) = \langle u(\rho), v(\sigma) \rangle$ for maps u and v into a Hilbert space $\mathcal{H}$ with $\|u(\rho)\| \leq R_u$ and $\|v(\sigma)\| \leq R_v$. Then for every $\delta > 0$, we have*

$$d_C(\delta) \leq 16 \left(\frac{R_u R_v}{\delta} \right)^2,$$

independently of the dimensions of H_1 and H_2 . In particular, we have $d_C(\delta) \leq 16/\delta^2$ for the swap operator F and for $[A = B]$, and $d_C(\delta) \leq 4/\delta^2$ for the swap test P_+ and the Hong–Ou–Mandel projection P_- .

Proof. The idea is to run the perceptron along the tree, always moving to the child that contradicts its prediction; the leaf reached at the end is a linear separator with margin $\delta/2$, and the perceptron makes few mistakes against such a separator. Regard the Hilbert space as a real inner product space with $\langle \cdot, \cdot \rangle_{\mathbb{R}} = \operatorname{Re} \langle \cdot, \cdot \rangle$; since φ_C is real, we have $\varphi_C = \langle u, v \rangle_{\mathbb{R}}$. Let a δ -tree of height n be given, with threshold γ and sign θ . The tree requires values of φ_C on both sides of γ , so we have $|\gamma| \leq R_u R_v$. Put $\hat{u}(\rho) = (u(\rho), R_u)$ and $\hat{v}(\sigma) = (v(\sigma), -\gamma/R_u)$ in $\mathcal{H} \oplus \mathbb{R}$; then $\langle \hat{u}(\rho), \hat{v}(\sigma) \rangle_{\mathbb{R}} = \varphi_C(\rho, \sigma) - \gamma$, $\|\hat{u}\| \leq \sqrt{2}R_u$ and $\|\hat{v}\| \leq \sqrt{2}R_v$. Start with $z_0 = 0$ at the root. At the node s reached at step i , let $\hat{y} = \operatorname{sign} \langle z_i, \hat{v}(\sigma_s) \rangle_{\mathbb{R}}$, with $\operatorname{sign} 0 = 1$, move to the child whose label $y = \theta(2i - 1)$ equals $-\hat{y}$, and put $z_{i+1} = z_i + y \hat{v}(\sigma_s)$. After n steps we reach a branch η , and its leaf $\hat{u}^* = \hat{u}(\rho_\eta)$ satisfies $y \langle \hat{u}^*, \hat{v}(\sigma_s) \rangle_{\mathbb{R}} \geq \delta/2$ at every step. Hence we have $\langle \hat{u}^*, z_n \rangle_{\mathbb{R}} \geq n\delta/2$, while $\|z_n\|^2 \leq \sum_i \|\hat{v}(\sigma_{s_i})\|^2 \leq 2nR_v^2$, because each step adds $2y \langle z_i, \hat{v}(\sigma_{s_i}) \rangle_{\mathbb{R}} \leq 0$ to the cross term. By the Cauchy–Schwarz inequality, we have $n\delta/2 \leq \sqrt{2}R_u \cdot \sqrt{2n}R_v$, that is, $n \leq 16(R_u R_v / \delta)^2$.

For F and $[A = B]$, we have $R_u, R_v \leq 1$ (see Section 4.3 and Proposition 4.6). Since $\varphi_{P_{\pm}} = \frac{1}{2}(1 \pm \varphi_F)$, every δ -tree for $P_{\pm}$ is a 2δ -tree for F . $\square$

Remark 9.12 (Node-dependent thresholds). The proof of Theorem 9.11 uses the threshold only through the probe vectors $\hat{v}(\sigma_s)$, and the bound $|\gamma| \leq R_u R_v$ holds at each node separately, since every node has leaves on both sides of its threshold. The proof therefore applies without change to trees in which the threshold γ_s depends on the node s . Consequently, the response class $\{\sigma \mapsto \varphi_C(\rho, \sigma) : \rho \in \mathfrak{D}(H_1)\}$ of a Hilbert-factorable observable has sequential fat-shattering dimension at most $16(R_u R_v / \delta)^2$ at scale δ in the sense of [RST15], and so it is online learnable in that sense, with bounds independent of the dimension.

9.4 Affine approximation versus exponential growth

The table collects the statements proved in this paper that correspond, for bipartite observables, to the dichotomy of deep computations described in the Introduction (see [ADIW24, DIMW⁺26]).

	stable side	unstable side
idealized responses	limits of <i>affine</i> combinations of actual responses, that is, of finite mixtures (see Theorem 6.1)	for the order projection, not approximable within any error less than 1 (see Example 6.3)
preparations resolved by n adaptive probes at margin δ	$O(n^{d_C(\delta)})$ if C is uniformly stable (see Theorem 9.9)	2^n at some margin (see Theorem 9.4)
Hilbert-factorable observables	tree dimension at most $16/\delta^2$, and $4/\delta^2$ for the swap test and HOM; mixtures of $\lceil 4/\varepsilon^2 \rceil$ preparations (see Theorems 9.11 and 6.5)	—

The approximating combinations in the first row are affine because expectations are affine in the state (see Remark 6.4); the polynomial in the second row has degree $d_C(\delta)$, which need not be one: for the comparison GT_n of Section 11, it is at least $n - 1$ at margin 1. In the Hilbert-factorable case, two further arguments are linear: idealized responses are inner products with a fixed vector, and tree heights are bounded by a linear classifier. The first two rows concern different notions. The sector comparison of Example 9.8 is stable, so it has affine simulation of idealized preparations, but it is not uniformly stable, so its growth at margin 1 is 2^n .

9.5 Learning: Littlestone's mistake bound

Decision trees of this kind also govern mistake bounds in online learning [Lit88]. Fix a threshold γ and a margin $\delta > 0$, and consider the following game. Alice prepares an unknown state ρ . In each round, Bob chooses a probe σ for which $|\varphi_C(\rho, \sigma) - \gamma| \geq \delta/2$; a learner predicts whether $\varphi_C(\rho, \sigma)$ lies above or below γ , and is then told the answer. Let $\text{mist}_C(\gamma, \delta)$ be the least integer m such that some deterministic learner makes at most m mistakes, for every ρ and every sequence of probes, and let $\text{mist}_C(\delta) = \sup_\gamma \text{mist}_C(\gamma, \delta)$. Let d_γ be the largest height of a δ -tree for C with threshold γ and sign $\theta = 1$. Since reflecting a tree exchanges $\theta = 1$ and $\theta = -1$, we have $\sup_\gamma d_\gamma = d_C(\delta)$.

Proposition 9.13 (Mistake bound equals tree dimension). *For every self-adjoint $C \in B(H_1 \otimes H_2)$, every γ and every $\delta > 0$, we have $\text{mist}_C(\gamma, \delta) = d_\gamma$. Consequently, the observable C is uniformly stable if and only if $\text{mist}_C(\delta)$ is finite for every $\delta > 0$, and then $\text{mist}_C(\delta) = d_C(\delta)$.*

Proof. If there is a tree of height n , an adversary walks down it, choosing at each node the child that contradicts the learner's prediction; the leaf reached is a preparation consistent with all answers, so every learner makes n mistakes. Conversely, run Littlestone's standard optimal algorithm: keep the set V of preparations consistent with the answers so far; for a probe σ , let V^1 and V^0 be the sets of elements of V above $\gamma + \delta/2$ and below $\gamma - \delta/2$, and predict the side ι for which $d(V^\iota)$ is larger, in the notation of Section 9.3. Arguing as in the proof of Theorem 9.9, the numbers $d(V^0)$ and $d(V^1)$ cannot both equal $d(V)$, so after a mistake the new consistent set $V^{1-\iota}$ satisfies $d(V^{1-\iota}) < d(V)$. Since the true preparation stays in V , we have $d(V) \geq 0$ throughout, and there are at most d_γ mistakes. The second statement follows from Theorem 9.6. $\square$

By Theorem 9.4, for an unstable observable the mistake bound in the threshold game is infinite at some margin, and the sector comparison of Example 9.8 shows that stability alone does not give a finite mistake bound. The threshold game is a restricted form of online learning: general online learnability of real-valued classes allows thresholds that vary from probe to probe (see the historical remarks below and Question (Q6)). In the Hilbert-factorable case the restriction is immaterial: by Remark 9.12, the perceptron of Theorem 9.11 learns online with at most $16(R_u R_v / \delta)^2$ mistakes even when the thresholds vary, in every dimension.

9.6 Historical remarks

Stability in a cardinal λ was introduced by Rowbottom [Row64] in his work on the Łoś conjecture for uncountable theories, extending Morley’s ω -stability [Mor65] to uncountable languages; Shelah studied the stability spectrum, the class of cardinals in which a theory is stable, and related it to the order property [She71, She90]. For Banach space structures, the stability spectrum theorem was proved by Iovino in his dissertation [Iov94]. Binary trees of formulas and the associated rank are due to Shelah [She71, She90], and the equivalence between bounded ladders, bounded trees and stability of a formula in all models of a theory, which underlies Theorem 9.6, is part of classical stability theory [She90]. For Banach spaces, the stability of all ultrapowers is Raynaud’s superstability [Ray83], the analytic counterpart of condition (3) of Theorem 9.6; Proposition 9.3 adapts [Iov, Proposition 2.3.1], where it is proved for first-order formulas, with the gap δ given by the failure of the double limit property in place of the negation of a formula, and with ultraproducts in place of compactness. Quantitative bounds in the converse direction come from a Ramsey argument for trees due to Hodges [Hod93]; Conant and Terry give explicit bounds for real-valued functions [CT26]. Theorem 9.9 is the tree analogue of the Sauer–Shelah lemma [Sau72, She72]; versions for Littlestone classes appear in [BDPSS09, RST15]. Littlestone introduced the dimension and the optimal algorithm that bear his name [Lit88]; for real-valued classes, online learnability is characterized by finiteness of the sequential fat-shattering dimension at every scale [RST15], which allows node-dependent thresholds (see Question (Q6)). The link between stability and Littlestone dimension in model theory is due to Chase and Freitag [CF19]. The bound in Theorem 9.11 is Novikoff’s mistake bound for the perceptron [Nov62, Blo62], read as a bound on tree height. For n -qubit states probed by two-outcome measurements, Aaronson, Chen, Hazan, Kale and Nayak showed that the sequential fat-shattering dimension is $O(n)$ at each fixed scale [ACH⁺18].

10 Order: stable observables are closed under norm limits and orthogonal sums

The goal of this section is to determine the algebraic and topological operations under which stability is preserved, and to locate the boundary of these closure properties: we show that stability survives norm limits and orthogonal sums, but not meets or strong limits.

10.1 Topological and algebraic closure

Proposition 10.1 (Topological and algebraic closure). *The set $\mathcal{O}$ of stable observables is a norm-closed operator system: a self-adjoint linear subspace of $B(H_1 \otimes H_2)$ that contains I and is closed in the norm topology. The set $\mathcal{O}$ contains the minimal tensor prod-*

uct $B(H_1) \otimes_{\min} B(H_2)$ and every observable whose expectation factors through a Hilbert space. The same holds for the set of uniformly stable observables.

Proof. The identity $I = I \otimes I$ is stable, and for density operators we have $\varphi_{C^*} = \overline{\varphi_C}$. Linear combinations and norm limits are handled by Lemma 3.4(a) and (b), and for uniform stability the same lemma applies to Φ_C^u . The last two statements are Theorem 4.1, Proposition 4.4 and Proposition 9.7. $\square$

Corollary 10.2 (Orthocomplements and orthogonal sums). *The stable projections form an orthomodular poset: if P is stable, then so is $I - P$, and if P and Q are stable with $PQ = 0$, then the projection $P + Q$ is stable.*

Proof. If the projections P and Q are stable, then the projections $I - P$, $P + Q$ (when $PQ = 0$) and $Q - P$ (when $P \leq Q$) are linear combinations of stable observables, and hence they are stable by Proposition 10.1. Thus the stable projections are closed under orthocomplements and orthogonal sums, and the orthomodular law, which holds for all projections on $H_1 \otimes H_2$, holds among them. $\square$

10.2 The boundary of closure: meets and strong limits

In Proposition 10.3 below, we show that the boundary of these closure properties lies at meets.

Proposition 10.3 (The boundary: meets). *If H_1 and H_2 are infinite-dimensional, then there are projections P_1 and P_2 in $B(H_1) \otimes_{\min} B(H_2)$, hence uniformly stable, whose meet $P_1 \wedge P_2$ is unstable.*

Proof. The idea is to hide the order projection inside the intersection of two subspaces that are compact perturbations of a product projection, by tilting each matrix entry into a fresh direction through a small angle, in opposite senses for the two subspaces. Let $(e_k) \cup (e'_k)$ be an orthonormal sequence in H_1 and $(f_l) \cup (f'_l)$ one in H_2 , and let $P_{k < l}$ be the order projection built from (e_k) and (f_l) . For $k \geq l$, let $\zeta_k = 1/k$ and

$$n_{kl} = \cos \zeta_k e_k \otimes f_l + \sin \zeta_k e'_k \otimes f'_l, \quad n'_{kl} = \cos \zeta_k e_k \otimes f_l - \sin \zeta_k e'_k \otimes f'_l,$$

and let P_1 and P_2 be the projections onto the closed subspaces $\text{ran } P_{k < l} \oplus \overline{\text{span}}\{n_{kl} : k \geq l\}$ and $\text{ran } P_{k < l} \oplus \overline{\text{span}}\{n'_{kl} : k \geq l\}$. The planes spanned by $e_k \otimes f_l$ and $e'_k \otimes f'_l$ are mutually orthogonal and orthogonal to the range of $P_{k < l}$, and in each of them the lines through n_{kl} and n'_{kl} meet only in 0, since $0 < \zeta_k < \pi/2$. Both subspaces are block diagonal for this decomposition, so their intersection is the range of $P_{k < l}$, that is, $P_1 \wedge P_2 = P_{k < l}$, which is unstable by Theorem 4.7.

Let Q_1 and Q_2 be the projections onto the closed spans of (e_k) and (f_l) . The operator $P_1 - Q_1 \otimes Q_2$ is the orthogonal sum over $k \geq l$ of the rank-two operators $|n_{kl}\rangle\langle n_{kl}| - |e_k \otimes f_l\rangle\langle e_k \otimes f_l|$, of norm $\sin \zeta_k$. For every $\varepsilon > 0$, only finitely many pairs $k \geq l$ have $1/k \geq \varepsilon$, so this sum is compact. Since $Q_1 \otimes Q_2$ and the compact operators lie in $B(H_1) \otimes_{\min} B(H_2)$, so does P_1 , and likewise P_2 . Conclude with Theorem 4.1 and Proposition 9.7. $\square$

By von Neumann's alternating projection theorem, the meet is the strong limit $P_1 \wedge P_2 = \lim_n (P_1 P_2)^n$ [Deu01]. The convergence is in norm if and only if the sum of the ranges of P_1 and P_2 is closed [Deu01]. In Proposition 10.3, the angles ζ_k accumulate at 0, so the sum of the ranges is not closed. Since $B(H_1) \otimes_{\min} B(H_2)$ is an algebra, the operators $(P_1 P_2)^n$ are stable, and their strong limit $P_1 \wedge P_2$ is unstable. Thus stability passes to norm limits (see Lemma 3.4(b)), but not to strong limits.

11 Outlook: registers of growing size

The goal of this section is to indicate how the uniform notions of Section 9 apply to quantum computation, where the natural setting is a family of registers $H_n = (\mathbb{C}^2)^{\otimes n}$ with readout observables C_n on $H_n \otimes H_n$, and deep computations idealized as ultralimits $\lim_{k, \mathcal{U}} \mathcal{E}_k \circ \cdots \circ \mathcal{E}_1(\rho_0)$ of outputs of quantum channels $\mathcal{E}_k$, in the sense of [ADIW24, DIMW⁺26]. Each C_n acts on a finite-dimensional space and is therefore stable (see Corollary 4.2); the meaningful condition is the uniform stability of the family, in the sense of the remark after Theorem 9.6.

Two families illustrate the distinction. The equality tests $\text{EQ}_n = \sum_x |x\rangle\langle x| \otimes |x\rangle\langle x|$ in the computational basis factor through ℓ^2 with constants independent of n (see the proof of Proposition 4.6). The family is therefore uniformly stable, its tree dimension is at most $16/\delta^2$ for all n (see Theorem 9.11), so its growth is polynomial of a degree independent of n , and the Monte Carlo bound of Corollary 6.6 holds with $M = \lceil 4/\varepsilon^2 \rceil$ for all n . The comparisons $\text{GT}_n = \sum_{x > y} |x\rangle\langle x| \otimes |y\rangle\langle y|$ contain 1-ladders of length $2^n - 1$, hence 1-trees of height $n - 1$ by Proposition 9.3: the family is not uniformly stable, and $n - 1$ adaptive probes resolve 2^{n-1} basis preparations. Let us notice that the same dichotomy appears in communication complexity: equality has constant public-coin randomized cost, while Greater-Than on n -bit inputs has cost $\Theta(\log n)$ [KN97, Vio15]. In the language of graphs, the communication matrix of Greater-Than is a *half-graph*, which is a ladder in the sense of Definition 9.1, and a family of communication problems excludes large Greater-Than instances exactly when its matrices contain no large half-graphs. Harms, Wild and Zamaraev showed that this exclusion, which is stability, is necessary for constant randomized cost [HWZ22], and Harms and Zamaraev showed that it is sufficient for matrices of sign-rank at most 3 and for unit disk graphs [HZ23].

12 Questions

We close with questions raised by the results above, on the structure of $\mathcal{O}$, on the gap between stability and uniform stability, and on quantitative bounds:

- (Q1) Is every stable observable a norm limit of observables whose slice maps factor through a Hilbert space?
- (Q2) Is a self-adjoint observable C uniformly stable if and only if T_C factors through a superreflexive Banach space?
- (Q3) Is the operator system $\mathcal{O}$ closed under products, and hence a C^* -algebra? If so, the meet of two stable projections whose ranges have a closed sum would be stable.
- (Q4) Is the meet of two *commuting* stable projections stable?
- (Q5) For a uniformly stable family of readouts on n -qubit registers, what is the minimal size of a mixture simulating idealized preparations to precision ε , and when is it independent of n ? Theorem 6.5 answers this for Hilbert-factorable families.
- (Q6) Theorem 9.9 and Proposition 9.13 concern trees with a constant threshold. Does uniform stability imply that the response class has finite sequential fat-shattering dimension at every scale, with node-dependent thresholds? Remark 9.12 gives a positive answer for Hilbert-factorable observables.

- (Q7) Is the tree dimension $d_C(\delta)$ of a uniformly stable observable always bounded by a polynomial in $1/\delta$, as in the Hilbert case, and is the degree of that polynomial related to the type of a space through which T_C factors?

References

- [ACH⁺18] S. Aaronson, X. Chen, E. Hazan, S. Kale, and A. Nayak, *Online learning of quantum states*, Advances in Neural Information Processing Systems 31 (NeurIPS 2018), 2018, pp. 8962–8972.
- [ADIW24] Samson Alva, E. Dueñez, J. N. Iovino, and C. Walton, *Approximability of deep computations*, 2024, Accepted; to appear in Mathematical Structures in Computer Science. arXiv:2409.06064.
- [Are51] R. Arens, *The adjoint of a bilinear operation*, Proc. Amer. Math. Soc. **2** (1951), 839–848.
- [BCWdW01] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, *Quantum fingerprinting*, Phys. Rev. Lett. **87** (2001), 167902.
- [BDPSS09] S. Ben-David, D. Pál, and S. Shalev-Shwartz, *Agnostic online learning*, Proc. 22nd Conference on Learning Theory (COLT 2009), 2009.
- [Bir35] G. Birkhoff, *Abstract linear dependence and lattices*, Amer. J. Math. **57** (1935), 800–804.
- [BL71] J. T. Baldwin and A. H. Lachlan, *On strongly minimal sets*, J. Symbolic Logic **36** (1971), 79–96.
- [Blo62] H. D. Block, *The perceptron: a model for brain functioning. I*, Rev. Modern Phys. **34** (1962), 123–135.
- [BY14] Itai Ben Yaacov, *Model theoretic stability and definability of types, after A. Grothendieck*, Bull. Symb. Log. **20** (2014), no. 4, 491–496. MR 3294276
- [BYBHU08] Itai Ben Yaacov, Alexander Berenstein, C. Ward Henson, and Alexander Usvyatsov, *Model theory for metric structures*, Model theory with applications to algebra and analysis. Vol. 2, London Math. Soc. Lecture Note Ser., vol. 350, Cambridge Univ. Press, Cambridge, 2008, pp. 315–427. MR MR2436146
- [BYU10] Itai Ben Yaacov and Alexander Usvyatsov, *Continuous first order logic and local stability*, Trans. Amer. Math. Soc. **362** (2010), no. 10, 5213–5259. MR 2657678
- [CF19] H. Chase and J. Freitag, *Model theory and machine learning*, Bull. Symbolic Logic **25** (2019), 319–332.
- [CR70] H. H. Crapo and G.-C. Rota, *On the Foundations of Combinatorial Theory: Combinatorial Geometries*, preliminary ed., MIT Press, Cambridge, MA, 1970.
- [CT26] G. Conant and C. Terry, *Encoding orders and trees in real-valued functions*, arXiv:2607.21761, 2026.
- [Deu01] F. Deutsch, *Best Approximation in Inner Product Spaces*, CMS Books in Mathematics, Springer, New York, 2001.
- [DFJP74] W. J. Davis, T. Figiel, W. B. Johnson, and A. Pelczynski, *Factoring weakly compact operators*, J. Functional Analysis **17** (1974), 311–327.

- [DIMW⁺26] E. Dueñez, J. Iovino, T. Matos-Wiederhold, L. Salvetti, and F. D. Tall, *Complexity of deep computations via topology of function spaces*, 2026, In preparation. arXiv:2601.00528.
- [Ebe47] W. F. Eberlein, *Weak compactness in Banach spaces. I*, Proc. Nat. Acad. Sci. U.S.A. **33** (1947), 51–53.
- [Gan40] V. Gantmacher, *Über schwache totalstetige Operatoren*, Mat. Sbornik N.S. **7(49)** (1940), 301–308.
- [Gro52] A. Grothendieck, *Critères de compacité dans les espaces fonctionnels généraux*, Amer. J. Math. **74** (1952), 168–186. MR 47313
- [Hei80] S. Heinrich, *Ultraproducts in Banach space theory*, J. Reine Angew. Math. **313** (1980), 72–104.
- [HH84] V. Harnik. and L. Harrington, *Fundamentals of forking*, Ann. Pure Appl. Logic **26** (1984), 245–286.
- [Hod93] Wilfrid Hodges, *Model theory*, Encyclopedia of Mathematics and its Applications, vol. 42, Cambridge University Press, Cambridge, 1993. MR 94e:03002
- [HOM87] C. K. Hong, Z. Y. Ou, and L. Mandel, *Measurement of subpicosecond time intervals between two photons by interference*, Phys. Rev. Lett. **59** (1987), 2044–2046.
- [HP11] E. Hrushovski and A. Pillay, *On NIP and invariant measures*, J. Eur. Math. Soc. **13** (2011), 1005–1061.
- [HPP08] Ehud Hrushovski, Ya’acov Peterzil, and Anand Pillay, *Groups, measures, and the NIP*, J. Amer. Math. Soc. **21** (2008), no. 2, 563–596. MR 2373360 (2008k:03078)
- [HWZ22] N. Harms, S. Wild, and V. Zamaraev, *Randomized communication and implicit graph representations*, STOC ’22: Proc. 54th Annual ACM Symposium on Theory of Computing, 2022.
- [HZ23] N. Harms and V. Zamaraev, *Randomized communication and implicit representations for matrices and graphs of small sign-rank*, arXiv:2307.04441, 2023.
- [Iov] J. Iovino, *Notes on stability theory*, preliminary notes, The University of Texas at San Antonio.
- [Iov94] José N. Iovino, *Stable theories in functional analysis*, Ph.D. thesis, University of Illinois at Urbana-Champaign, 1994.
- [Iov99a] ———, *Stable Banach spaces and Banach space structures. I. Fundamentals*, Models, algebras, and proofs (Bogotá, 1995), Dekker, New York, 1999, pp. 77–95. MR 2000h:03064a
- [Iov99b] ———, *Stable models and reflexive Banach spaces*, J. Symbolic Logic **64** (1999), no. 4, 1595–1600. MR 2001e:03068
- [Kei87] H. J. Keisler, *Measures and forking*, Ann. Pure Appl. Logic **34** (1987), 119–169.
- [KM81] Jean-Louis Krivine and Bernard Maurey, *Espaces de Banach stables*, Israel J. Math. **39** (1981), no. 4, 273–295. MR 83a:46030
- [KN97] E. Kushilevitz and N. Nisan, *Communication Complexity*, Cambridge University Press, 1997.

- [KP97] Byunghan Kim and Anand Pillay, *Simple theories*, Ann. Pure Appl. Logic **88** (1997), no. 2-3, 149–164, Joint AILA-KGS Model Theory Meeting (Florence, 1995). MR MR1600895 (99b:03049)
- [KR83] R. V. Kadison and J. R. Ringrose, *Fundamentals of the Theory of Operator Algebras, Vol. I*, Academic Press, 1983.
- [KŠ40] M. Krein and V. Šmulian, *On regularly convex sets in the space conjugate to a Banach space*, Ann. of Math. **41** (1940), 556–583.
- [Lev07] E. Levy, *Weakly compact “matrices”, Fubini-like property and extension of densely defined semigroups of operators*, arXiv:0704.3558, 2007.
- [Lit88] N. Littlestone, *Learning quickly when irrelevant attributes abound: a new linear-threshold algorithm*, Machine Learning **2** (1988), 285–318.
- [LP79] D. Lascar and B. Poizat, *An introduction to forking*, J. Symbolic Logic **44** (1979), 330–350.
- [Mak84] M. Makkai, *A survey of basic stability theory, with particular emphasis on orthogonality and regular types*, Israel J. Math. **49** (1984), 181–238.
- [Maz33] S. Mazur, *Über konvexe Mengen in linearen normierten Räumen*, Studia Math. **4** (1933), 70–84.
- [ML36] S. Mac Lane, *Some interpretations of abstract linear dependence in terms of projective geometry*, Amer. J. Math. **58** (1936), 236–240.
- [Mor65] M. Morley, *Categoricity in power*, Trans. Amer. Math. Soc. **114** (1965), 514–538.
- [NK09] H. Nishimura and S. Kuroda (eds.), *A Lost Mathematician, Takeo Nakasawa: The Forgotten Father of Matroid Theory*, Birkhäuser, Basel, 2009.
- [Nov62] A. B. J. Novikoff, *On convergence proofs on perceptrons*, Proc. Symposium on the Mathematical Theory of Automata, vol. 12, Polytechnic Institute of Brooklyn, 1962, pp. 615–622.
- [Pil96] A. Pillay, *Geometric stability theory*, The Clarendon Press Oxford University Press, New York, 1996, Oxford Science Publications.
- [Pis81] G. Pisier, *Remarques sur un résultat non publié de B. Maurey*, Séminaire d’Analyse Fonctionnelle 1980–1981, Exp. V, École Polytechnique, Palaiseau, 1981.
- [Ptá63] V. Pták, *A combinatorial lemma on the existence of convex means and its application to weak compactness*, Convexity, Proc. Sympos. Pure Math., vol. 7, Amer. Math. Soc., 1963, pp. 437–450.
- [Pym65] J. S. Pym, *The convolution of functionals on spaces of bounded functions*, Proc. London Math. Soc. (3) **15** (1965), 84–104.
- [Ray83] Y. Raynaud, *Espaces de Banach superstables, distances stables et homéomorphismes uniformes*, Israel J. Math. **44** (1983), no. 1, 33–52. MR 84j:46024
- [Row64] F. Rowbottom, *The $\aleph_1$ conjecture for uncountable theories*, Notices of the Amer. Math. Soc. **11** (1964), 248–249.
- [RST15] A. Rakhlin, K. Sridharan, and A. Tewari, *Online learning via sequential complexities*, J. Mach. Learn. Res. **16** (2015), 155–186.
- [Sau72] N. Sauer, *On the density of families of sets*, J. Combinatorial Theory Ser. A **13** (1972), 145–147. MR 307902

- [She71] Saharon Shelah, *Stability, the f.c.p., and superstability; model theoretic properties of formulas in first order theory*, Ann. Math. Logic **3** (1971), no. 3, 271–362. MR 0317926
- [She72] ———, *A combinatorial problem; stability and order for models and theories in infinitary languages*, Pacific J. Math. **41** (1972), 247–261. MR 307903
- [She90] S. Shelah, *Classification theory and the number of nonisomorphic models*, second ed., Studies in Logic and the Foundations of Mathematics, vol. 92, North-Holland Publishing Co., Amsterdam, 1990. MR MR1083551 (91k:03085)
- [Sim68] S. Simons, *The iterated limit condition, a Fubini theorem and weak compactness*, Math. Ann. **176** (1968), 87–95.
- [Ste10] E. Steinitz, *Algebraische Theorie der Körper*, J. Reine Angew. Math. **137** (1910), 167–309.
- [Ü87] A. Ülger, *Weakly compact bilinear forms and Arens regularity*, Proc. Amer. Math. Soc. **101** (1987), 697–704.
- [vdW30] B. L. van der Waerden, *Moderne Algebra*, Springer, Berlin, 1930.
- [Vio15] E. Viola, *The communication complexity of addition*, Combinatorica **35** (2015), 703–747.
- [Š40] V. Šmulian, *Über lineare topologische Räume*, Mat. Sbornik N.S. **7(49)** (1940), 425–448.
- [Whi35] H. Whitney, *On the abstract properties of linear dependence*, Amer. J. Math. **57** (1935), 509–533.