

SHELAH’S DIVIDING LINES IN QUANTUM MEASUREMENT: 32 VARIATIONS IN THREE PARTS

JOSE’ IOVINO

In homage to Alexander Grothendieck, Saharon Shelah and Stevo Todor’ević

ABSTRACT. Logicians and operator theorists have long used the same instrument, the ultrapower, without suspecting that they were describing the same things; this paper is an attempt to introduce them to each other. The logician will discover that the dividing lines which Shelah drew in order to classify theories serve equally well to classify quantum measurements. Stability turns out to be nothing more mysterious than the indifference of a measurement to the order in which two idealized limits are taken; NIP sorts physical orders, from the counting of photons to the causal order of spacetime, into the trichotomy of Todor’ević; and simplicity, the most abstract of the three, is found at work in quantum error correction, where forking separates the errors that do no harm from those that do. The operator theorist will discover what types and independence are, and that both notions have been within reach all along. Both will find a single theorem of Grothendieck at the bottom of everything. A reader who belongs to neither community may read it also, provided that ultrafilters hold no terrors, being less alarming than their name.

CONTENTS

Aria	5
Variatio 1. a 1 Clav. · <i>Observables as structures</i>	7
Part I. Stability	10
Variatio 2. a 1 Clav. · <i>The order property</i>	10
Variatio 3. a 1 Clav. Canone all’Unisuono · <i>Formulas and observables</i>	12
Variatio 4. a 1 Clav. · <i>Hilbert space</i>	14
Variatio 5. a 1 ô vero 2 Clav. · <i>Stability in seven languages</i>	15
Variatio 6. a 1 Clav. Canone alla Seconda · <i>The independent product</i>	18
Variatio 7. a 1 ô vero 2 Clav. Al tempo di Giga · <i>Discrimination and learning</i>	19
Variatio 8. a 2 Clav. · <i>Bounds independent of the dimension</i>	21
Variatio 9. a 1 Clav. Canone alla Terza · <i>Mixtures simulate idealized preparations</i>	22
Variatio 10. a 1 Clav. Fughetta · <i>Countable randomization does not detect instability</i>	24
Part II. NIP	25
Variatio 11. a 2 Clav. · <i>The independence property</i>	25
Variatio 12. Canone alla Quarta · <i>The rational order</i>	26
Variatio 13. a 2 Clav. · <i>The dictionary of NIP</i>	28
Variatio 14. a 2 Clav. · <i>The Todor’ević trichotomy</i>	30
Variatio 15. a 1 Clav. Canone alla Quinta. Andante · <i>The heptachotomy</i>	32
Variatio 16. a 1 Clav. Ouverture · <i>Physical orders in every class</i>	35

Date: Working draft, October 14, 2026.

Variatio 17.	a 2 Clav. · <i>Orders of limits in phase transitions</i>	37
Variatio 18.	a 1 Clav. Canone alla Sesta · <i>A quantum ferromagnet in the classification</i>	38
Variatio 19.	a 1 Clav. · <i>Convexity forces metrizability</i>	40
Part III. Simplicity		41
Variatio 20.	a 2 Clav. · <i>Simplicity</i>	41
Variatio 21.	Canone alla Settima · <i>The Rado graph</i>	42
Variatio 22.	a 1 Clav. Alla breve · <i>The Pauli group</i>	44
Variatio 23.	a 2 Clav. · <i>The dictionary of simplicity</i>	46
Variatio 24.	a 1 Clav. Canone all'Ottava · <i>Fraïssé limits and Clifford circuits</i>	47
Variatio 25.	a 2 Clav. Adagio · <i>Forking and logical errors</i>	48
Variatio 26.	a 2 Clav. · <i>The gluing of syndromes</i>	51
Variatio 27.	a 2 Clav. Canone alla Nona · <i>The Rado graph inside the Pauli group</i>	52
Variatio 28.	a 2 Clav. · <i>Mixing destroys simplicity</i>	53
Variatio 29.	a 1 ô vero 2 Clav. · <i>Actual and idealized trees</i>	55
Variatio 30.	a 1 Clav. Quodlibet · <i>The commutation relation</i>	56
Aria da capo e fine		57
Questions		58
Historical remarks		58
Acknowledgments		59
References		60

The operator theorist and the logician possess, without having agreed upon it, a common instrument: the ultraproduct. To the first it is a construction, by which a Banach space or an operator algebra is enlarged until bounded sequences acquire limits; to the second it is an elementary extension, the most concrete there is, in which every element has a type over the structure from which it came. The difference is not one of fact but of what each regards as worth noticing. The logician notices types, and the independence of one type from another, and on these two notions Shelah erected his classification of theories. The operator theorist, with the same ultrapowers at hand, has had less occasion to use them. Conversely, the physics in which these notions acquire their meaning is seldom set before the logician in a form that a logician would accept as a definition. This paper undertakes both tasks: it explains types and independence to the operator theorist, and it states the physics and the quantum information for the logician, formally and completely.

The thesis is that, in this setting, the abstractions of logic become instruments of physical description. If, in place of the formulas of a first-order theory, one considers bipartite observables, then the dividing lines that Shelah drew among theories become distinctions among measurements, and the theorems by which each line is characterized become assertions about what can be observed. The results are those of a series written with Eduardo Dueñez [DI26a, DI26b, DI26c]; the present account gathers them into one.

The Aria states the double limit criterion of Grothendieck and its physical meaning; at the end the Aria returns, da capo, and is heard against what the three parts have established. Between them lie thirty sections, divided into three parts of identical design, in order that each may be read against the others. Part I treats stability, Part II NIP, Part III simplicity. Each part begins with the configuration that its line forbids, the theorem that characterizes the line, and a Fraïssé limit in which the line is exhibited: Hilbert space for stability, the rational

order for NIP, the random graph and the Pauli group for simplicity. A dictionary then renders the model theory of the line in physical terms. Each part ends with mixing, the operation that distinguishes quantum observables from arbitrary formulas: a state may be mixed, the response of a bipartite observable is affine in each state, and every observable therefore carries with it the mixtures of its preparations. Many sections close with a paragraph for the operator theorist, one for the logician, and a sentence recording what has been established. Table 1 sets the three parts side by side; the box below states the main results.

	I. Stability	II. NIP	III. Simplicity
forbidden configuration	a ladder (order property)	an independent sequence	a tree whose siblings are k -inconsistent, for one fixed k
characterization	double limits commute (Grothendieck); an independence relation with stationarity over models (Harnik, Harrington)	Baire class one (Rosenthal; Bourgain, Fremlin, Talagrand)	an independence relation with the independence theorem over models (Kim, Pillay)
nonforking extensions of a type over a model	unique	invariant over the model, possibly several	possibly several, amalgamated by the independence theorem
Fraïssé limit	Hilbert space	the rational order	the random graph; the Pauli group
dictionary	orders of idealization	physical orders	quantum error correction
what mixing does	realizes idealized preparations	makes hereditary separability imply metrizable	destroys simplicity

TABLE 1. The three parts side by side.

Main results.

- (I) *Stability* [DI26a]. An observable is stable exactly when the two orders of idealization agree. This condition has seven equivalent forms, each with an operational meaning, among them simulation of idealized preparations by finite mixtures and the existence of a unique independent product; in a uniform form, it is equivalent to polynomial growth of the number of preparations that adaptive probes can tell apart (see Theorems 5.1 and 7.1).
- (II) *NIP* [DI26b]. An observable is NIP exactly when its idealized responses are of the first Baire class. For observables, hereditary separability of the idealized responses implies metrizability, so the trichotomy of Todorčević has three classes, each containing a physical order: photon numbers, positions, and the causal order of spacetime, joined in the last class by the comparison of phases on a circle. Each of the seven minimal families of the heptachotomy of Argyros, Dodos and Kanellopoulos is realized by a tree of preparations of photons, of a particle in a box, or of a spin chain (see Theorems 11.1, 19.1, 15.1 and 16.1).
- (III) *Simplicity* [DI26c]. The Pauli group on countably many qubits is the Fraïssé limit of finite commutation patterns, and the notions of simple theories translate into quantum error correction. Mixing preparations destroys simplicity: a diagonal observable whose pure responses have the independence property is not simple, because its mixtures carry the tree property of the second kind (see Variations 23–26 and Theorem 28.1).

Figure 1 places the observables of Parts I and II in their classes.

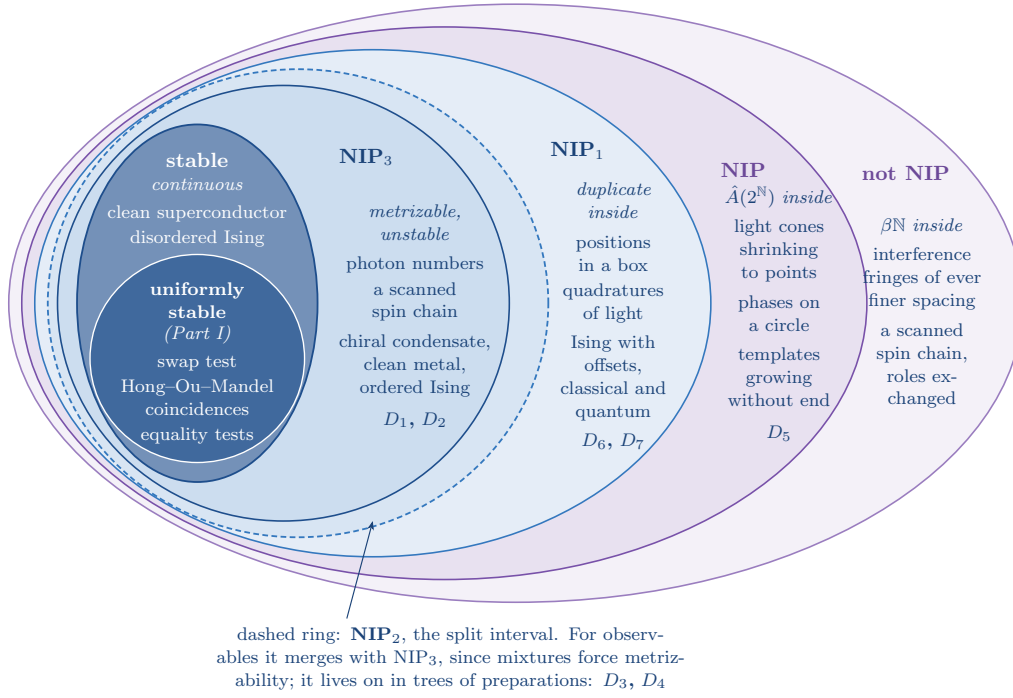


FIGURE 1. The classes of stable and NIP observables, from the innermost (continuous idealized responses) outward, with the physical observables that inhabit them and the minimal families $D_1, \dots, D_7$ that their trees of preparations realize.

ARIA

Physics lives on idealized limits. A physicist speaks freely of systems that no laboratory prepares: a particle with a sharp momentum, a magnet in infinite volume, an oscillator excited to infinite energy. To a mathematician each of them is an ultrafilter limit: given actual preparations, density operators ρ_n , and a nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$, the idealized system is the state $\omega(A) = \lim_{n, \mathcal{U}} \text{Tr}(\rho_n A)$ of $B(H)$, which need not be given by any density operator. In the language of model theory, once an observable is fixed, ω determines the type of an element of an ultrapower, restricted to the formula that the observable defines (see Variatio 1). A sequence of energy eigenstates whose energy grows without bound has no limit among density operators, but it has ultrafilter limits among the states of $B(H)$, and these formalize the idealized oscillator. For a single system, idealization is harmless: every prediction about the idealized system is the limit of the predictions about the actual ones.

The difficulty begins with two systems. Quantum information describes its protocols in terms of parties: two experimenters, by custom Alice and Bob, each of whom controls one system in a separate laboratory. Alice's system is described by a Hilbert space H_1 and Bob's by a Hilbert space H_2 ; each party prepares and measures states of its own system, and neither acts directly on the other's. The two systems together are described by the tensor product $H_1 \otimes H_2$, and a *bipartite observable* is a bounded self-adjoint operator C on it: a measurement performed on the pair, whose outcome depends jointly on both systems. When Alice prepares ρ and Bob, independently, prepares σ , the joint state is $\rho \otimes \sigma$, and the expected outcome of C is $\text{Tr}((\rho \otimes \sigma)C)$. The two roles are symmetric in principle; this paper speaks of Alice's preparations and Bob's probes because Alice's state is the object under study and Bob's

states are the tests applied to it. A logician will recognize the division: a bipartite observable is a formula $\varphi(x; y)$ whose variables are partitioned, as Shelah partitions them, with Alice holding x and Bob holding y , and $\text{Tr}((\rho \otimes \sigma)C)$ is its value at the pair (ρ, σ) . When both systems are idealized, the prediction becomes one of two double limits,

$$\lim_n \lim_m \text{Tr}((\rho_n \otimes \sigma_m)C) \quad \text{or} \quad \lim_m \lim_n \text{Tr}((\rho_n \otimes \sigma_m)C),$$

according to whose system is idealized first, and the two need not agree. An oscillator is the Hilbert space $L^2(\mathbb{R})$ with the energy observable, or *Hamiltonian*, $\frac{1}{2}(-d^2/dq^2 + q^2)$, an unbounded self-adjoint operator; its eigenvectors $e_0, e_1, \dots$ form an orthonormal basis, and in the state e_k a measurement of energy gives $k + \frac{1}{2}$ with certainty. Only the basis and the order of the energies are used below, through bounded spectral projections. Suppose that two identical oscillators are prepared in energy eigenstates e_k and f_l , and that the measurement asks whether the second has the higher energy. The answer a_{kl} is 1 when $k < l$, and 0 when $k \geq l$. Then we have

$$\lim_l \lim_k a_{kl} = 0, \quad \lim_k \lim_l a_{kl} = 1.$$

The matrix (a_{kl}) has ones above the diagonal and zeros on and below it; we call this configuration the *staircase*. Once both energies are idealized, the event “Bob’s oscillator has the higher energy” has no well-defined probability; its value depends on which system was idealized first.

The agreement of the two orders determines the operational meaning of an idealized system. When they agree, every idealized preparation of Alice is simulated, to any accuracy and uniformly over Bob’s probes, by finite mixtures of actual preparations; two idealized systems have a unique independent product; and, in the uniform version, the number of preparations that n adaptive measurements can tell apart grows only polynomially in n , so that an unknown preparation can be learned. When they disagree, all of this fails, and at some margin n measurements tell apart 2^n preparations for every n (see Theorems 5.1 and 7.1).

The theme of this paper is the theorem that says exactly when the two orders agree. Grothendieck proved it in 1952, in his work on compactness in function spaces [Gro52]: let K be a compact space, and let A be a bounded set of continuous functions on K . Then every function in the pointwise closure of A is continuous if and only if, for all sequences (f_n) in A and (x_m) in K ,

$$\lim_n \lim_m f_n(x_m) = \lim_m \lim_n f_n(x_m)$$

whenever both iterated limits exist. The first condition says that A is relatively weakly compact in the Banach space $C(K)$; a topological property of the whole family is decided by its countable grids. Let Alice’s preparations be the functions and Bob’s probes the points. Then the theorem says that a bipartite measurement is insensitive to the order of idealization exactly when the idealized responses of Alice are continuous on the compact set of Bob’s states, idealized or not.

A logician recognizes the staircase as the order property, the configuration that defines the first of the dividing lines that Shelah drew across the first-order theories; read for a formula, the criterion of Grothendieck is the definability of types in a stable theory [BY14]. The second line, NIP, separates the failures of the double limit in which idealized responses remain limits of sequences of actual ones from those in which they do not (see Theorem 11.1), and the third, simplicity, replaces limits by the amalgamation of independent information.

Part I studies the observables that satisfy the criterion of Grothendieck, Part II the ways in which it can fail while idealized responses remain sequential limits of actual ones, and Part III the amalgamation of independent information that takes its place in simple theories.

Variatio 1. A 1 CLAV.

Observables as structures

The staircase involves two idealized systems at once, and so far its double limits have been computed only along sequences of eigenstates. In this variation we define idealized states as states of $B(H)$ and the two products that correspond to the two orders of idealization. It then explains, for operator theorists, how ultrapowers lead to the model-theoretic notions of elementary extension, type and independence of types, of which idealized states and their products are instances.

States and observables. This subsection is addressed to logicians. A quantum system is described by a complex separable Hilbert space H . A unit vector $\psi \in H$ describes a *pure* preparation; a general preparation is described by a *density operator*, a positive trace-class operator ρ of trace 1, and the pure preparations correspond to the rank-one projections $|\psi\rangle\langle\psi|$; here $|\psi\rangle\langle\chi|$ denotes the operator $\xi \mapsto \langle\xi, \chi\rangle\psi$, and a unit vector ψ also stands for the state $|\psi\rangle\langle\psi|$. The density operators form a convex set $\mathfrak{D}(H)$, and a *mixture* $\sum_i t_i \rho_i$, with $t_i \geq 0$ and $\sum_i t_i = 1$, describes a preparation that produces ρ_i with probability t_i . A measurement with finitely many outcomes is a family of projections $E_1, \dots, E_r$ with sum I , and in the state ρ the outcome j occurs with probability $\text{Tr}(\rho E_j)$, by the rule of Born. More generally, a bounded self-adjoint operator A on H is an *observable*; by the spectral theorem it is a norm limit of real linear combinations of commuting projections, and $\text{Tr}(\rho A)$ is the expected value of its outcome in the state ρ . The expectation is affine in ρ , and this affineness is the formal content of the statement that mixing is probabilistic.

Two systems, Alice's and Bob's, are described together by the Hilbert tensor product $H_1 \otimes H_2$. When they are prepared independently in states ρ and σ , the joint state is $\rho \otimes \sigma$, and a bounded self-adjoint operator C on $H_1 \otimes H_2$, a *bipartite observable*, has expectation

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

Once Alice has prepared ρ , the observable acts on Bob's system alone as the *effective observable* $T_C(\rho)$, the unique bounded operator on H_2 with $\text{Tr}(\sigma T_C(\rho)) = \varphi_C(\rho, \sigma)$ for every density operator σ ; it is the partial trace $T_C(\rho) = \text{Tr}_1((\rho \otimes I)C)$. The function $\sigma \mapsto \varphi_C(\rho, \sigma)$ is the *response* of Alice's preparation to Bob's probes. For the energy comparison, with energy eigenstates e_k and f_l , the observable is the projection

$$P = \sum_{k < l} |e_k\rangle\langle e_k| \otimes |f_l\rangle\langle f_l|, \quad T_P(e_k) = \sum_{l > k} |f_l\rangle\langle f_l|.$$

Idealized states. The sequence (e_k) has no limit among density operators. Its limit points are *idealized states*: positive linear functionals ω on the algebra $B(H)$ of bounded operators with $\omega(I) = 1$. They form a compact convex set in the weak* topology, in which the density operators are dense. For a nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$, the state $\omega(A) = \lim_{k, \mathcal{U}} \langle A e_k, e_k \rangle$ vanishes on every compact operator, so it assigns probability zero to every finite set of levels: it is the oscillator excited to infinite energy.

Two orders of idealization. Let ω_1 and ω_2 be states of $B(H_1)$ and $B(H_2)$, and let (ρ_i) and (σ_j) be nets of density operators converging to them in the weak* topology. The expectation of C can be idealized in two orders:

$$(\omega_1 \otimes_{\mathbf{A}} \omega_2)(C) = \lim_j \lim_i \varphi_C(\rho_i, \sigma_j), \quad (\omega_1 \otimes_{\mathbf{B}} \omega_2)(C) = \lim_i \lim_j \varphi_C(\rho_i, \sigma_j),$$

with Alice idealized first or Bob idealized first. Let us notice that both limits exist and do not depend on the nets: for the first, $\sigma \mapsto \lim_i \varphi_C(\rho_i, \sigma)$ is a bounded linear functional on trace-class operators, hence of the form $\sigma \mapsto \text{Tr}(\sigma Z)$ for a unique operator $Z = Z_{\omega_1}$ on H_2 ; then we have $(\omega_1 \otimes_A \omega_2)(C) = \omega_2(Z_{\omega_1})$. The second is symmetric. The two products are the two Arens extensions of the bilinear form φ_C to the biduals [Are51].

Let us notice that if one of the two states is actual, the two products agree: if ω_1 is given by a density operator ρ , then we have $Z_{\omega_1} = T_C(\rho)$, and both products equal $\omega_2(T_C(\rho))$; the case of ω_2 is symmetric. Disagreement needs two idealized systems.

For the energy comparison, take $\omega_1 = \lim_{k, \mathcal{U}} e_k$ and $\omega_2 = \lim_{l, \mathcal{U}} f_l$. For each fixed l , the value $\varphi_P(e_k, f_l)$ is 0 for all $k \geq l$; for each fixed k , it is 1 for all $l > k$. Hence we have

$$(\omega_1 \otimes_A \omega_2)(P) = 0, \quad (\omega_1 \otimes_B \omega_2)(P) = 1.$$

The two orders of idealization assign to the event “Bob’s energy is higher” the probabilities 0 and 1.

Ultrapowers, elementary extensions and types. This subsection is addressed to operator theorists. In the continuous logic of Ben Yaacov, Berenstein, Henson and Usvyatsov [BYBHU08], a structure consists of complete metric spaces, its *sorts*, together with uniformly continuous real-valued functions on finite products of the sorts, its *predicates*; the metric of each sort is a distinguished predicate, and all predicates take values in a fixed bounded interval. In the real-valued logic of Keisler [Kei23], which is the setting of this paper, a *general structure* consists of a nonempty universe for each sort and an interpretation of each symbol of its vocabulary; each predicate symbol P takes its values in a bounded interval $[r_P, s_P]$ that depends on P . A general structure determines a vocabulary but no metric signature: no distinguished distance predicate and no moduli of uniform continuity are prescribed. When the vocabulary is countable, Keisler’s Expansion Theorem shows that every theory has a pre-metric expansion, obtained by adding a distance predicate that is a uniform limit of formulas, and that this expansion is unique up to uniform equivalence. A bipartite observable C gives a general structure $\mathcal{M}_C$ with two sorts, the state spaces $\mathfrak{D}(H_1)$ and $\mathfrak{D}(H_2)$, and three predicates: the expectation φ_C , which takes its values in $[-\|C\|, \|C\|]$, and, on each sort, the trace distance $(\rho, \rho') \mapsto \|\rho - \rho'\|_1$, which takes its values in $[0, 2]$. Formulas are built from the predicates by continuous functions $u: \mathbb{R}^n \rightarrow \mathbb{R}$ and by the quantifiers $\sup_x$ and $\inf_x$; they take real values rather than the values true and false.

Let $\mathcal{U}$ be a nonprincipal ultrafilter on $\mathbb{N}$. The ultrapower $\mathcal{M}_C^{\mathcal{U}}$ has as sorts the powers $\mathfrak{D}(H_1)^{\mathbb{N}}$ and $\mathfrak{D}(H_2)^{\mathbb{N}}$ modulo Leibniz equality. Since the trace distance is a predicate and the expectation is Lipschitz in each state, two sequences (ρ_n) and (ρ'_n) are identified exactly when $\lim_{n, \mathcal{U}} \|\rho_n - \rho'_n\|_1 = 0$, so the sorts are the metric ultrapowers of the state spaces. Its predicates are the ultralimits of those of $\mathcal{M}_C$; for instance, $\varphi([\rho_n], [\sigma_n]) = \lim_{n, \mathcal{U}} \varphi_C(\rho_n, \sigma_n)$. This is the ultrapower of Banach space theory [Hei80], applied to the state spaces. Operator theorists take ultrapowers on both sides of the duality between states and observables. The algebra $B(H)$ is ultrapowered in the operator norm, and the von Neumann algebra generated by its Banach ultrapower is the ultrapower of Groh and Raynaud [Gro84, Ray02]; for tracial von Neumann algebras one uses instead the 2-norm of the trace, and for σ -finite ones the strong* topology of Ocneanu (see [AH14]). The states form the predual, which is ultrapowered in the trace norm, and Raynaud showed that the Banach ultrapower of the predual is the predual of the ultrapower of Groh and Raynaud [Ray02]. The trace norm is also the metric that measurements see: the largest difference between the probabilities that a single measurement assigns to an outcome under ρ and under ρ' is $\frac{1}{2}\|\rho - \rho'\|_1$, so the sorts of $\mathcal{M}_C^{\mathcal{U}}$ identify exactly

the sequences of preparations that no sequence of measurements tells apart in the limit. The operator norm would identify too much: if P_n and Q_n are orthogonal projections of rank n , then the operator norm of $P_n/n - Q_n/n$ is $1/n$, whereas its trace norm is 2, and the projections P_n tell the two preparations apart with certainty. The theorem of Łoś, in its continuous form, states that for every formula $\theta(x_1, \dots, x_k)$ and all elements $[a_n^1], \dots, [a_n^k]$ of the ultrapower, we have

$$\theta^{\mathcal{M}_C^{\mathcal{U}}}([a_n^1], \dots, [a_n^k]) = \lim_{n, \mathcal{U}} \theta^{\mathcal{M}_C}(a_n^1, \dots, a_n^k).$$

In particular, the diagonal map, which sends a to the class of the constant sequence, preserves the value of every formula: it is an *elementary embedding*. In general, a structure $\mathcal{N}$ containing $\mathcal{M}$ is an *elementary extension* of $\mathcal{M}$ if every formula with parameters from $\mathcal{M}$ takes the same value in $\mathcal{N}$ as in $\mathcal{M}$. The ultrapower is the most concrete elementary extension, and it is enough for Parts I and II: since $\mathcal{U}$ is countably incomplete and the language is countable, $\mathcal{M}_C^{\mathcal{U}}$ is $\aleph_1$ -saturated, and, since $\mathcal{M}_C$ is separable, every separable elementary extension of $\mathcal{M}_C$ embeds elementarily in $\mathcal{M}_C^{\mathcal{U}}$ over $\mathcal{M}_C$ [BYBHU08]. Every state of $B(H)$ is $\lim_{n, \mathcal{U}} \rho_n$ for some sequence of density operators and some ultrafilter $\mathcal{U}$ on $\mathbb{N}$, since the density operators form a norm-separable, weak*-dense subset of the states; thus no idealized state is lost by using sequences. Part III needs a larger elementary extension, the monster model of Variatio 20.

Elementary extensions are where types live. Let $\mathcal{N}$ be an elementary extension of $\mathcal{M}$, and let a be an element of $\mathcal{N}$. The *type* of a over $\mathcal{M}$ is the function that assigns to every formula $\theta(x, y)$ and every tuple b from $\mathcal{M}$ the value $\theta^{\mathcal{N}}(a, b)$; it records everything that formulas with parameters from $\mathcal{M}$ say about a . Its restriction to the single formula $\varphi(x, y)$ is the φ -*type* of a . For an element $a = [\rho_n]$ of Alice's sort in $\mathcal{M}_C^{\mathcal{U}}$, the φ -type over $\mathcal{M}_C$ is the function

$$\sigma \longmapsto \lim_{n, \mathcal{U}} \text{Tr}((\rho_n \otimes \sigma)C)$$

on Bob's density operators, that is, the response of the idealized state $\omega = \lim_{n, \mathcal{U}} \rho_n$, the limit in the weak* topology. Thus an idealized state, read through one observable, determines a φ -type; different idealized states may have the same φ -type, and the full type of $[\rho_n]$ is not a function of ω , since it also records, for instance, trace-norm distances. Every type over $\mathcal{M}_C$ is realized in $\mathcal{M}_C^{\mathcal{U}}$.

Independence enters with two systems. Let p and q be types over $\mathcal{M}_C$ in Alice's and in Bob's sort, realized by $a = [\rho_n]_{\mathcal{U}}$ and by $b = [\sigma_m]_{\mathcal{V}}$. To evaluate φ at the pair, the two elements must be placed in a common elementary extension. Placing both in the same ultrapower gives the joint limit $\lim_{n, \mathcal{U}} \varphi_C(\rho_n, \sigma_n)$, which in general is neither iterated limit and does not make a and b independent; for the staircase with $\rho_n = e_n$ and $\sigma_n = f_n$ it is 0. An iterated ultrapower offers two other ways. If a is realized in $\mathcal{M}_C^{\mathcal{U}}$ and b then in $(\mathcal{M}_C^{\mathcal{U}})^{\mathcal{V}}$, the predicate takes the value $\lim_{m, \mathcal{V}} \lim_{n, \mathcal{U}} \varphi_C(\rho_n, \sigma_m)$; in the other order it takes the value $\lim_{n, \mathcal{U}} \lim_{m, \mathcal{V}} \varphi_C(\rho_n, \sigma_m)$. These are the two orders of idealization. In the first, the type of b over $\mathcal{M}_C^{\mathcal{U}}$ is the $\mathcal{V}$ -limit of the types over $\mathcal{M}_C^{\mathcal{U}}$ of the elements σ_m of $\mathcal{M}_C$. A type over a set $B \supseteq \mathcal{M}$ with this property, that of being approximately finitely satisfiable in $\mathcal{M}$, is a *coheir* of its restriction to $\mathcal{M}$. The *theory* of a structure assigns to each formula without free variables, a *sentence*, its value; elementary extensions have the same theory, and a *model* of a theory is a structure with that theory. A formula is *stable in a theory* if it has no infinite ladder (see Variatio 2) in any model of the theory. In Variatio 6 below, the two orders of the iterated ultrapower become the subject of independence of types.

For the operator theorist. The ultrapower $\mathcal{M}_C^{\mathcal{U}}$ is built from the metric ultrapowers of the state spaces, as in Banach space theory, and the theorem of Łoś makes it an elementary extension of $\mathcal{M}_C$: the diagonal embedding preserves the value of every formula, not only that of the predicate φ_C . The φ -type over $\mathcal{M}_C$ of an element $[\rho_n]$ is the response of the weak* limit state $\lim_{n,\mathcal{U}} \rho_n$, while its full type records more, such as trace-norm distances. The two orders of realization in an iterated ultrapower give the two products $\otimes_A$ and $\otimes_B$, and in Variatio 6 we show that they are the two independent amalgams of two types.

For the logician. A preparation of a quantum system is a density operator ρ , and a measurement with outcomes $1, \dots, r$ is a family of projections $E_1, \dots, E_r$ with sum I , the outcome j occurring with probability $\text{Tr}(\rho E_j)$; the number $\varphi_C(\rho, \sigma)$ is the expected outcome of the observable C when Alice's system is prepared in ρ and Bob's, independently, in σ . An idealized state is a state of $B(H)$ that need not be given by a density operator: the state $\lim_{k,\mathcal{U}} e_k$ of the oscillator vanishes on every compact operator, so it gives probability zero to every finite set of energy levels, and no laboratory prepares it. The two orders of idealization are the two Arens extensions of the bilinear form φ_C , and for the energy comparison they assign to the event “Bob's energy is higher” the probabilities 0 and 1. What distinguishes observables from arbitrary formulas is that the expectation is affine in each state, and states can be mixed.

In Theorem 3.1 below, we show that the double limit condition of Grothendieck, read for the expectation of C , is the statement that the two products agree for every pair of states. The staircase of the oscillators fails it with the largest possible gap.

Part I. Stability

Variatio 2. A 1 CLAV.

The order property

Two atoms with discrete spectra are compared by two measurements. The first asks whether they are in levels of the same energy, the second whether Bob's energy is the higher. If A and B are the two Hamiltonians, the energy observables of the two atoms, with spectral projections Π_λ^A and Π_λ^B , the two measurements are the projections

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

Both are diagonal in a product basis of eigenvectors, so both lie in a commutative algebra; in this sense they are classical. Nevertheless, in Proposition 2.2 below we show that they behave in opposite ways once the atoms are excited without bound. In this variation we define ladders and stability, show that an observable is unstable exactly when it has an infinite ladder, and explain the opposite behavior of the two tests.

Ladders. Let C be a bipartite observable with expectation $\varphi_C(\rho, \sigma) = \text{Tr}((\rho \otimes \sigma)C)$, and let $\delta > 0$ be given. A *ladder at margin δ* consists of preparations $\rho_1, \rho_2, \dots$ of Alice, probes $\sigma_1, \sigma_2, \dots$ of Bob and a threshold t such that

$$\varphi_C(\rho_i, \sigma_j) \leq t - \frac{\delta}{2} \text{ for } i \leq j, \quad \varphi_C(\rho_i, \sigma_j) \geq t + \frac{\delta}{2} \text{ for } i > j,$$

or the same with the two inequalities exchanged. The ladder may be finite or infinite. A formula with ladders of every finite length has the *order property* of Shelah [She71]. The observable C is *stable* if it satisfies the double limit condition of Grothendieck: for all sequences

(ρ_k) and (σ_l) ,

$$\lim_l \lim_k \varphi_C(\rho_k, \sigma_l) = \lim_k \lim_l \varphi_C(\rho_k, \sigma_l)$$

whenever both iterated limits exist.

Figure 2 draws the staircase of the introduction as a ladder.

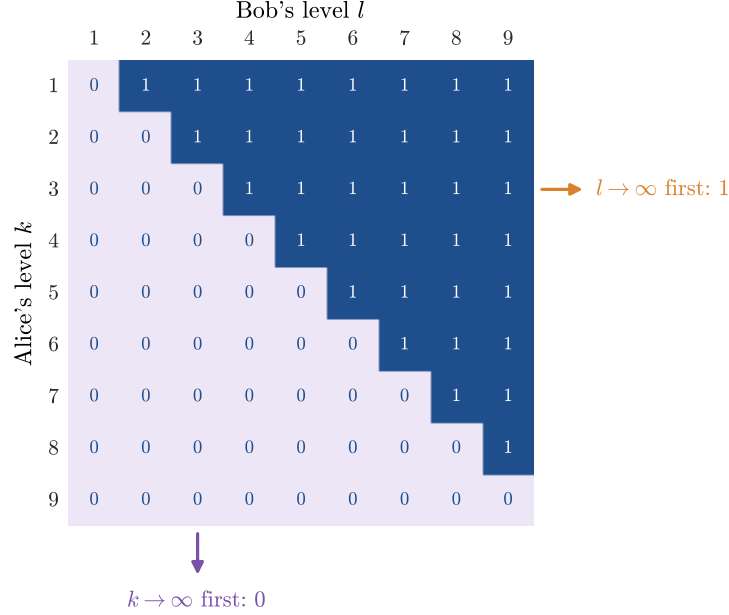


FIGURE 2. The energy comparison of two oscillators: the value $\varphi(e_k, f_l)$ is 1 when Bob's level l exceeds Alice's level k and 0 otherwise. Every row tends to 1 as $l \rightarrow \infty$, and every column tends to 0 as $k \rightarrow \infty$, so the two iterated limits are 1 and 0.

Theorem 2.1. *An observable is unstable if and only if it has an infinite ladder at some margin.*

Proof. Given an infinite ladder, pass to subsequences of rows and of columns along which both iterated limits exist; they lie on opposite sides of t , at distance at least δ . Conversely, let $a_{kl} = \varphi_C(\rho_k, \sigma_l)$ have iterated limits $L_1 = \lim_k \lim_l a_{kl}$ and $L_2 = \lim_l \lim_k a_{kl}$ with $L_1 \neq L_2$, and put $\alpha_k = \lim_l a_{kl}$ and $\beta_l = \lim_k a_{kl}$, so that $\alpha_k \rightarrow L_1$ and $\beta_l \rightarrow L_2$. Fix $\varepsilon > 0$ and choose indices alternately. Choose $k_m > k_{m-1}$ with $|\alpha_{k_m} - L_1| < \varepsilon$ and $|a_{k_m l_p} - \beta_{l_p}| < \varepsilon$ for all $p < m$, which holds for all large k . Then choose $l_m > l_{m-1}$ with $|\beta_{l_m} - L_2| < \varepsilon$ and $|a_{k_p l_m} - \alpha_{k_p}| < \varepsilon$ for all $p \leq m$. For $p \leq m$ the value $a_{k_p l_m}$ is within 2ε of L_1 , and for $p > m$ it is within 2ε of L_2 . With $4\varepsilon < |L_1 - L_2|$, this is an infinite ladder at margin $|L_1 - L_2| - 4\varepsilon$. $\square$

Proposition 2.2. *The equality test $[A = B]$ is stable. If the spectra of A and B are both unbounded above, the order test $[A < B]$ is unstable.*

Proof. For the equality test, $\varphi(\rho, \sigma) = \sum_\lambda \text{Tr}(\rho \Pi_\lambda^A) \text{Tr}(\sigma \Pi_\lambda^B) = \langle u(\rho), v(\sigma) \rangle$, an inner product in ℓ^2 of vectors of norm at most 1. Bounded sequences in a Hilbert space have weakly convergent subsequences; if $u(\rho_k) \rightarrow \xi$ and $v(\sigma_l) \rightarrow \xi'$ weakly, both iterated limits equal $\langle \xi, \xi' \rangle$. For the order test, choose eigenvalues $\lambda_1 < \lambda'_1 < \lambda_2 < \lambda'_2 < \dots$ with λ_k an eigenvalue of A and λ'_l one of B , and eigenvectors ψ_k and χ_l . Then the value $\varphi(\psi_i, \chi_j)$ is 1 for $i \leq j$

and 0 for $i > j$: an infinite ladder at margin 1, with the two inequalities of the definition exchanged. $\square$

The instability of $[A < B]$ is the classical failure of the double limit for the order of $\mathbb{N}$, carried into a Hilbert space by two orthonormal sequences; it owes nothing to noncommutativity. The equality test is stable for the opposite reason: it is an inner product, and the inner product has no infinite ladders.

For the operator theorist. A ladder is the order property of Shelah [She71], in the real-valued form of Krivine and Maurey [KM81] and of Iovino [Iov99a]. Ultrapowers turn finite ladders into infinite ones: if for every n there is a ladder $(\rho_i^n, \sigma_j^n)_{i,j \leq n}$ at margin δ with threshold t_n , then, by the theorem of Łoś, the elements $a_i = [\rho_i^n]_{\mathcal{U}}$ and $b_j = [\sigma_j^n]_{\mathcal{U}}$, with arbitrary entries for $n < \max(i, j)$, form an infinite ladder at margin δ and threshold $\lim_{n, \mathcal{U}} t_n$ in $\mathcal{M}_C^{\mathcal{U}}$. Hence the formula φ_C is stable in the theory of $\mathcal{M}_C$, in the sense of Variatio 1, exactly when C is uniformly stable, that is, when its ladders have bounded length at every margin (see Variatio 7). Stability of C is weaker: it forbids infinite ladders in $\mathcal{M}_C$ only, and Theorem 2.1 identifies it with the double limit condition.

For the logician. The Hamiltonian of an atom is the self-adjoint operator whose spectral projection Π_λ projects onto the states of energy λ ; in such a state, a measurement of energy returns λ with certainty. The tests $[A = B]$ and $[A < B]$ are yes-no measurements on the pair of atoms, and $\varphi(\rho, \sigma)$ is the probability of the answer yes when Alice's atom is prepared in ρ and Bob's in σ . For the equality test this probability is $\sum_\lambda p_\lambda q_\lambda$, where p_λ and q_λ are the probabilities that the two atoms have energy λ , an inner product of two vectors of ℓ^2 ; for the order test, eigenstates of increasing energy, taken alternately from Alice's atom and from Bob's, give answers that are certain and switch at the diagonal. Both tests commute with every energy measurement on either atom, so the instability of $[A < B]$ comes from interleaved sequences of energy levels and not from noncommutativity.

A ladder is the combinatorial configuration behind a failure of the double limit: two disagreeing iterated limits yield an infinite ladder, and the criterion of Grothendieck is exactly the absence of infinite ladders.

Variatio 3. A 1 CLAV. CANONE ALL'UNISUONO

Formulas and observables

In this variation we set out the dictionary of stability (see Table 2). Each theorem is stated once and read twice, by the logician and by the physicist, and the same words carry a different meaning in each reading.

The logician. Let $\varphi(x, y)$ be a formula, and let p be a type in x , the limit of elements a_i along an ultrafilter. The type is *definable* if the value $\varphi(p, b) = \lim_i \varphi(a_i, b)$ is given by a function of b that extends continuously to all types in y . The formula is stable exactly when every such type is definable [BY14]. In the language of Variatio 1, let $a = [\rho_n]_{\mathcal{U}}$, let b be an element of Bob's sort in an elementary extension $\mathcal{N}$ of $\mathcal{M}_C$, and let a' be the class of (ρ_n) in $\mathcal{N}^{\mathcal{U}}$, which realizes the coheir of $\text{tp}(a/\mathcal{M}_C)$ over $\mathcal{N}$. The type of a is definable exactly when $\varphi(a', b)$, the value with Bob idealized first, depends continuously on $\text{tp}(b/\mathcal{M}_C)$.

The physicist. Let C be a bipartite observable, and let ω_1 be an idealized state of Alice, the limit of preparations ρ_i . Its response $\sigma \mapsto \lim_i \varphi_C(\rho_i, \sigma)$ is given by an operator: there is a unique Z_{ω_1} on Bob's space with $\lim_i \varphi_C(\rho_i, \sigma) = \text{Tr}(\sigma Z_{\omega_1})$ for every density operator σ . The response is *itself an observable* if the same operator also gives the value at idealized

The logician	The physicist
a structure with two sorts	Alice's and Bob's systems
an element of a sort	an actual preparation, a density operator
a formula $\varphi(x, y)$	the expectation $\varphi_C(\rho, \sigma)$ of an observable
an elementary extension, such as an ultrapower	the state spaces enlarged by idealized states
the φ -type of an element of an ultrapower	the response of an idealized state, the ultrafilter limit of preparations
the φ -type of a , as a function of b	the response of a preparation to the probes
a definable type	an idealized response that is itself an observable
realizing one coheir before the other	idealizing one system before the other
an independent amalgam of two types	a product of two idealized states
the order property	a ladder
a stable formula	an observable whose two orders of idealization agree

TABLE 2. The dictionary of stability.

probes, in the order with Bob idealized first: we have $(\omega_1 \otimes_B \omega_2)(C) = \omega_2(Z_{\omega_1})$ for every state ω_2 . In Theorem 3.1 below, we show that the observable is stable exactly when every idealized response is itself an observable.

Theorem 3.1 (Grothendieck [Gro52]; [DI26a, Theorem 5.1]). *For a bipartite observable C , the following are equivalent:*

- (1) *the observable C is stable;*
- (2) *for every state ω_1 of Alice and every state ω_2 of Bob, we have $(\omega_1 \otimes_A \omega_2)(C) = (\omega_1 \otimes_B \omega_2)(C)$;*
- (3) *every idealized response of C is itself an observable.*

Proof. One argument serves both readings. Since $(\omega_1 \otimes_A \omega_2)(C) = \omega_2(Z_{\omega_1})$ always, conditions (2) and (3) say the same thing. If (2) fails, take nets (ρ_i) and (σ_j) converging to ω_1 and ω_2 , with iterated limits $L_1 = \lim_i \lim_j$ and $L_2 = \lim_j \lim_i$ that differ. Put $\alpha_i = \lim_j \varphi_C(\rho_i, \sigma_j)$ and $\beta_j = \lim_i \varphi_C(\rho_i, \sigma_j)$, fix $\varepsilon > 0$ with $4\varepsilon < |L_1 - L_2|$, and choose indices alternately: i_m with α_{i_m} within ε of L_1 and $\varphi_C(\rho_{i_m}, \sigma_{j_p})$ within ε of β_{j_p} for $p < m$; then j_m with β_{j_m} within ε of L_2 and $\varphi_C(\rho_{i_p}, \sigma_{j_m})$ within ε of α_{i_p} for $p \leq m$. Each step asks for finitely many conditions that hold eventually along a net. The value $\varphi_C(\rho_{i_p}, \sigma_{j_m})$ is then within 2ε of L_1 for $p \leq m$ and of L_2 for $p > m$, so along subsequences of (ρ_{i_p}) and (σ_{j_m}) on which both iterated limits exist, they differ, and C is unstable. Conversely, if C is unstable, there are sequences with iterated limits $L_1 = \lim_k \lim_l \neq L_2 = \lim_l \lim_k$; their limits along a nonprincipal ultrafilter are states ω_1 and ω_2 with $(\omega_1 \otimes_B \omega_2)(C) = L_1$ and $(\omega_1 \otimes_A \omega_2)(C) = L_2$. $\square$

The logician. In the logician's reading, the definition of the type ω_1 is the operator Z_{ω_1} , evaluated at types of the other sort. Theorem 3.1 is Grothendieck's theorem read as definability of types, as observed by Ben Yaacov [BY14].

The physicist. For the oscillators of the energy comparison P , take $\omega_1 = \lim_{k,u} e_k$. The effective observables $T_P(e_k) = \sum_{l>k} |f_l\rangle\langle f_l|$ decrease to 0, so we have $Z_{\omega_1} = 0$: on every actual probe, the idealized response of Alice takes the value 0. On the idealized probe $\omega_2 = \lim_{l,u} f_l$, with Bob idealized first, it takes the value 1, while $\omega_2(Z_{\omega_1}) = 0$. No operator reproduces this response.

Both readings. Grothendieck's condition, that the iterated limits agree, is the condition that idealized responses are again observables, that is, that types are definable. The type of the infinitely excited oscillator of the staircase is not definable, and its response is not an observable.

Variatio 4. A 1 CLAV.

Hilbert space

A *qubit* is the Hilbert space $\mathbb{C}^2$, with orthonormal basis $|0\rangle, |1\rangle$. In the *swap test*, two registers with the same Hilbert space, prepared in ρ and σ , are coupled to an auxiliary qubit, the *ancilla*, prepared in $(|0\rangle + |1\rangle)/\sqrt{2}$. The unitary that exchanges the two registers when the ancilla is in $|1\rangle$ and acts as the identity when it is in $|0\rangle$ is applied, and the ancilla is then measured in the basis $(|0\rangle \pm |1\rangle)/\sqrt{2}$, the outcome $\pm$ counting as acceptance. This is the swap test [BCWdW01], and it accepts with probability $\frac{1}{2}(1 + \text{Tr } \rho\sigma)$. In optics the same quantity appears without an ancilla. A photon is described by $\mathbb{C}^2 \otimes K$, where the factor $\mathbb{C}^2$ records which of two paths, or *ports*, it occupies and K its internal degrees of freedom, such as polarization or spatial mode; a balanced beam splitter acts on the port factor by the unitary $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, and two photons, being identical bosons, have their joint state in the symmetric subspace of $(\mathbb{C}^2 \otimes K)^{\otimes 2}$. Two photons whose internal states are ρ and σ enter the two ports of a balanced beam splitter, and they leave by different ports with probability $\frac{1}{2}(1 - \text{Tr } \rho\sigma)$, which vanishes for equal pure states. This is the effect of Hong, Ou and Mandel [HOM87]. Both experiments measure the overlap $\text{Tr } \rho\sigma$, and the overlap is an inner product. In this variation we show that every observable whose expectation is an inner product is stable, and that Hilbert space is the Fraïssé limit behind stability.

Lemma 4.1 (The swap trick). *Let F be the unitary on $K \otimes K$ with $F(\psi \otimes \chi) = \chi \otimes \psi$. Then we have $\text{Tr}((\rho \otimes \sigma)F) = \text{Tr}(\rho\sigma)$, the Hilbert–Schmidt inner product of ρ and σ . The swap test and the coincidence of Hong, Ou and Mandel are the projections $\frac{1}{2}(I \pm F)$.*

Proof. For $\rho = |\psi\rangle\langle\psi|$ and $\sigma = |\chi\rangle\langle\chi|$, the operator $(\rho \otimes \sigma)F$ is $|\psi \otimes \chi\rangle\langle\chi' \otimes \psi'|$, whose trace is $\langle\psi, \chi'\rangle\langle\chi, \psi'\rangle = \text{Tr}(\rho\sigma)$; extend by linearity and continuity. $\square$

Theorem 4.2. *If $\varphi_C(\rho, \sigma) = \langle u(\rho), v(\sigma) \rangle$ for bounded maps u and v into a Hilbert space, then the observable C is stable. In particular, the swap, the swap test and the coincidence of Hong, Ou and Mandel are stable.*

Proof. Given sequences along which both iterated limits exist, pass to subsequences with $u(\rho_k) \rightarrow \xi$ and $v(\sigma_l) \rightarrow \xi'$ weakly. For each k , the inner limit over l is $\langle u(\rho_k), \xi' \rangle$, which tends to $\langle \xi, \xi' \rangle$; the other order gives the same value. $\square$

Let us notice that if K is infinite-dimensional, the swap is not a norm limit of finite sums of products $A \otimes B$: such a limit has a compact effective map $\rho \mapsto T_C(\rho)$, while for the swap $T_F(\rho) = \rho$, and orthogonal pure states are sent to operators at distance 1. Thus the

stability of the swap does not come from approximation by local observables; it comes from Theorem 4.2, since the expectation is an inner product.

The Fraïssé limit. By the theorem of Fraïssé, a class of finite structures closed under substructures, with joint embedding and amalgamation, has a unique countable homogeneous limit that contains all its members [Fra54]; Ben Yaacov extended the theory to metric structures [BY15]. The finite-dimensional Hilbert spaces form such a class, since two isometric embeddings of a space E amalgamate in $E_1 \oplus_E E_2$, where the orthogonal complements of E are orthogonal to each other. The limit is the separable infinite-dimensional Hilbert space, and homogeneity is the extension of isometries between finite-dimensional subspaces to unitary operators.

For the operator theorist. The ultrapower $H^{\mathcal{U}}$ is again a Hilbert space and an elementary extension of H . For the inner product, the φ -type of $[\xi_n]$ over H is $y \mapsto \langle \xi, y \rangle$, with ξ the weak limit of (ξ_n) along $\mathcal{U}$: it is given by an element of H , which is definability of types (see Variatio 3), and by the theorem of Łoś the ladders of overlap measurements have bounded length at every margin. Over a closed subspace E , vectors a and b are independent in the sense of nonforking (see Variatio 6) exactly when $a - P_E a$ and $b - P_E b$ are orthogonal [BYBHU08]: the amalgam $E_1 \oplus_E E_2$ is the independent one.

For the logician. Both experiments measure the two-outcome observable whose outcomes are the projections $\frac{1}{2}(I \pm F)$. For pure states ψ and χ , the overlap $\text{Tr } \rho\sigma = |\langle \psi, \chi \rangle|^2$ is the probability that a system prepared in ψ passes a test for χ , so the experiments compare two unknown states without measuring either. In infinite dimension, the stability of the swap comes from Theorem 4.2, not from approximation by measurements on the two systems separately.

For an inner product, the weak limits of Alice's and of Bob's vectors are again vectors of the Hilbert space, and both iterated limits equal their inner product. Idealization never leaves the Hilbert space, and this is why Hilbert space is the Fraïssé limit associated with stability.

Variatio 5. A 1 Ô VERO 2 CLAV.

Stability in seven languages

The double limit condition has been found, under different names, in analysis, Banach space theory, operator algebras, probability and model theory. For a bipartite observable each of these forms says something about idealized preparations. The theorem below collects seven of them, and then one example is read in all seven. Condition (A2) is Theorem 3.1 of Variatio 3, and three of the conditions describe operations of a laboratory, developed in the variations that follow: the independent product (A7) in Variatio 6, the simulation of idealized preparations by finite mixtures (A5) in Variatio 9, and randomization (A6) in Variatio 10. Between them, Variations 7 and 8 treat the uniform version of the condition, which the ultrapower turns into a statement about every elementary extension.

Throughout, C is a bounded observable on $H_1 \otimes H_2$ with expectation $\varphi_C(\rho, \sigma) = \text{Tr}((\rho \otimes \sigma)C)$, the effective observable of a preparation ρ is $T_C(\rho)$, and $\otimes_A$ and $\otimes_B$ are the two orders of idealization, with Alice or Bob idealized first.

Theorem 5.1 ([DI26a, Theorem A]). *For every bounded observable C on $H_1 \otimes H_2$, the following are equivalent:*

- (A1) *the observable C is stable: the two iterated limits agree;*
- (A2) (Grothendieck) *the effective observables $T_C(\rho)$, for $\rho \in \mathfrak{D}(H_1)$, form a relatively weakly compact set, so that every idealized response is itself given by an observable;*
- (A3) (Iovino; Davis, Figiel, Johnson, Pełczyński) *the expectation is the pairing of a reflexive Banach space with its dual;*
- (A4) (Arens) *the two orders of idealization agree on all pairs of idealized states;*
- (A5) (Mazur) *every idealized preparation of Alice is simulated, to any accuracy and uniformly over all probes, by finite mixtures of actual preparations, with Bob's system idealized first;*
- (A6) (Pták, Simons) *iterated integrals of the expectation against any two finitely additive probability measures on the state spaces commute;*
- (A7) (Harrington) *for every pair of idealized preparations, the two orders of realizing the pair agree, and the common value is the unique, symmetric independent product.*

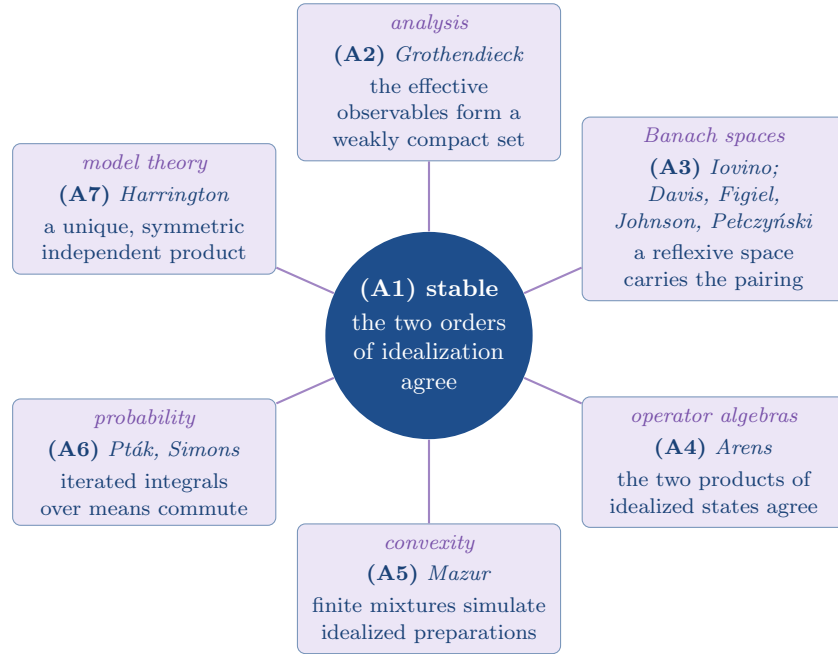


FIGURE 3. Seven readings of one condition. Each form of stability in Theorem 5.1 comes from a different field and says something different about idealized preparations.

The equivalence of the seven conditions. The equivalence of (A1) and (A4) is the extraction of a ladder from two disagreeing iterated limits. For (A2), the operator Z_{ω_1} that represents Alice's idealized response on actual probes is a weak* cluster point of the effective observables; under (A4) it is a weak cluster point, and the theorem of Eberlein and Šmulian turns weak cluster points into relative weak compactness [Ebe47]. A weakly compact operator factors through a reflexive space by the theorem of Davis, Figiel, Johnson

and Pelczyński [DFJP74], which gives (A3); Iovino proved that stable functions are exactly the restrictions of such pairings [Iov99b], and reflexivity yields (A1) because bounded sets of a reflexive space are weakly compact. For (A5), a convex set has the same weak and norm closures by the theorem of Mazur [Maz33], so a weak approximation of the idealized response by mixtures becomes a norm approximation, uniform over probes. For (A6), an iterated integral of an affine function depends only on the barycenters of the two measures, which are states, so the two iterated integrals are the two products of (A4); this is the theorem of Pták and Simons [Ptá63, Sim68]. For (A7), the two orders of realizing a pair of types are the two iterated ultrafilter limits, and their agreement is Harrington's lemma [HH84].

One example in seven languages. Let P be the energy comparison of two oscillators, with $\varphi_P(e_k, f_l) = 1$ for $k < l$ and 0 otherwise, and let ω_1 and ω_2 be the limits of (e_k) and (f_l) along a nonprincipal ultrafilter.

- (A1) The iterated limits are 0 and 1.
- (A2) The effective observables $\sum_{l>k} |f_l\rangle\langle f_l|$ decrease to 0, but the state ω_2 gives every one of them the value 1 and their limit the value 0; no subsequence converges weakly.
- (A3) No reflexive space carries the comparison: in a reflexive space the weak limits of Alice's and Bob's vectors would give both iterated limits the same value.
- (A4) The two products of ω_1 and ω_2 are 0 and 1.
- (A5) No mixture ρ of Alice's preparations simulates ω_1 within any error less than 1: the response of ω_1 is 0 on every f_l , while $\text{Tr}(f_l T_P(\rho)) = \sum_{k<l} \langle e_k, \rho e_k \rangle$ tends to 1 as l grows.
- (A6) For the means given by the ultrafilter on the two sequences, one iterated integral is 0 and the other is 1.
- (A7) The pair (ω_1, ω_2) is realized in two orders that assign the comparison the values 0 and 1; there is no independent product.

For the operator theorist. Read in the structure $\mathcal{M}_C$, condition (A2) says that every φ -type realized in the ultrapower is definable: its values at the elements of Bob's sort of any elementary extension, with Bob idealized first, are given by the operator Z_{ω_1} , a continuous function of their types (see Variatio 3). Condition (A6) concerns *Keisler measures*, the averaged analogue of types: a Keisler measure over $\mathcal{M}$ in the variable x is a regular Borel probability measure on the compact space of types in x over $\mathcal{M}$, equivalently a positive normalized linear functional on the formulas $\theta(x, b)$ with b in $\mathcal{M}$, and types are the Dirac measures. Ultrapowers produce them as they produce types: for finitely supported probability measures μ_n on Alice's sort, the functional $\theta(x, b) \mapsto \lim_{n, \mathcal{U}} \int \theta(x, b) d\mu_n(x)$ is a Keisler measure, and condition (A6) says that the two iterated integrals of φ_C against two such measures agree. Condition (A7) is Harrington's lemma (see Theorem 6.1): the two independent amalgams obtained in an iterated ultrapower coincide.

For the logician. Three of the seven conditions describe operations of a laboratory. A finite mixture $\sum_i t_i \rho_i$ is prepared by choosing the preparation ρ_i at random with probability t_i and forgetting the choice, and condition (A5) says that, for a stable observable, such a procedure reproduces an idealized preparation within ε , simultaneously for all of Bob's probes, in the order with Bob idealized first. Condition (A6) concerns preparations drawn at random from infinitely many (see Variatio 10), and condition (A7) concerns the joint state of two systems prepared independently, which is $\rho \otimes \sigma$ for actual preparations and must be defined by a limit for idealized ones. For the energy comparison all three fail with the largest possible gap,

because the event “Bob’s oscillator is more excited” has probability 0 or 1 according to the order in which the two infinitely excited oscillators are approached.

The seven conditions are seven readings of the double limit criterion of Grothendieck, and the staircase fails each of them in the same way, with the values 0 and 1.

Variatio 6. A 1 CLAV. CANONE ALLA SECONDA

The independent product

Two systems are prepared independently. If the preparations are actual, the joint state is $\rho \otimes \sigma$. If they are idealized, Variatio 1 gives two candidates, obtained by realizing the two types in an iterated ultrapower in either order. In this variation we state when the two candidates coincide, in two readings: the logician speaks of two-valued types over a model, the physicist of states of an algebra of observables.

We call a joint type of (a, b) obtained in either order of Variatio 1 an *independent amalgam* of p and q over $\mathcal{M}_C$; in general it depends on the ultrafilters, and not only on p and q . Whether independent amalgams are unique is the question that separates stable theories from unstable ones. If φ is stable in the theory, then over a model the coheir of a φ -type does not depend on the ultrafilter, it is the unique nonforking extension (forking is recalled below), and independence is symmetric [HH84]. For a single structure the relevant condition is weaker: in Theorem 6.2 below we show that the two orders agree on all pairs of states exactly when C is stable, that is, when φ_C has no infinite ladder in $\mathcal{M}_C$ itself, while stability of φ_C in the theory of $\mathcal{M}_C$ is uniform stability (see Theorem 7.1). In physical terms, an independent amalgam of two idealized states gives a candidate value for their product at C , and Theorem 6.2 assembles these values, over all stable observables, into a state. Simple theories retain a symmetric independence relation without uniqueness, and Part III shows what it becomes in quantum error correction.

The logician. Let M be a model of a stable theory, let p and q be complete types over M in variables x and y , and let $\varphi(x, y)$ be a formula. Stability gives definitions: there are formulas $d_p\varphi(y)$ and $d_q\varphi(x)$ over M such that $\varphi(x, b) \in p$ exactly when $M \models d_p\varphi(b)$, and symmetrically for q .

Theorem 6.1 (Harrington’s lemma [HH84]). *In a stable theory, $d_p\varphi(y) \in q$ if and only if $d_q\varphi(x) \in p$.*

Every type over a model is finitely satisfiable in it, so p is the trace of an ultrafilter $\mathcal{U}$ on the tuples of M , and q the trace of an ultrafilter $\mathcal{V}$. If φ is regarded as a $\{0, 1\}$ -valued function on pairs of tuples, then the two sides of the lemma read:

$$\lim_{b, \mathcal{V}} \lim_{a, \mathcal{U}} \varphi(a, b) = 1 \quad \text{and} \quad \lim_{a, \mathcal{U}} \lim_{b, \mathcal{V}} \varphi(a, b) = 1.$$

Harrington’s lemma is the exchange of two iterated ultrafilter limits. To state its role we recall forking. A formula $\theta(x, b)$ *divides* over a set A if there are $b = b_0, b_1, b_2, \dots$ in an elementary extension, indiscernible over A (any two increasing n -tuples have the same type over A), such that for some k no k of the formulas $\theta(x, b_i)$ have a common solution; in real-valued logic, k -inconsistency is required up to some $\varepsilon > 0$. A type *forks* over A if it implies a finite disjunction of formulas that divide over A , and $a \downarrow_A b$ means that $\text{tp}(a/Ab)$ does not fork over A . In every theory a coheir does not fork; hence the two orders of the iterated ultrapower of Variatio 1 always produce nonforking amalgams, and in a stable theory they produce the only one. Harrington’s lemma is the step that makes nonforking independence symmetric in a stable theory, and with uniqueness of nonforking extensions over models it gives the characterization of stable theories by Harnik and Harrington [HH84].

The physicist. Let $\mathcal{O}$ be the set of stable observables on $H_1 \otimes H_2$, a norm-closed self-adjoint subspace containing I [DI26a, Proposition 10.1]. For states ω_1 and ω_2 , let $\omega_1 \otimes \omega_2$ be the common value of the two orders of idealization on $\mathcal{O}$.

Theorem 6.2 ([DI26a, Theorem 8.2]). *For all states ω_1 of $B(H_1)$ and ω_2 of $B(H_2)$:*

- (a) (Existence) *the functional $\omega_1 \otimes \omega_2$ is a state of $\mathcal{O}$: linear, positive, with value 1 at I ;*
- (b) (Extension) *if ω_1 and ω_2 are density operators ρ and σ , then we have $(\omega_1 \otimes \omega_2)(C) = \text{Tr}((\rho \otimes \sigma)C)$;*
- (c) (Uniqueness) *for each $C \in \mathcal{O}$, the map $(\omega_1, \omega_2) \mapsto (\omega_1 \otimes \omega_2)(C)$ is the only function on pairs of states that is weak* continuous in each variable separately and satisfies (b);*
- (d) (Symmetry) *exchanging the roles of the two systems exchanges the two factors.*

An observable is stable if and only if the two orders of idealization assign it the same value for all pairs of states.

Idea of the proof. Positivity holds for each order separately, since each is an iterated limit of expectations of a positive operator in product states. Extension holds because the two orders agree when one factor is actual. For uniqueness, the product with Alice idealized first is weak* continuous in Bob's state, being evaluation at the operator Z_{ω_1} , and the product with Bob idealized first is continuous in Alice's; when they agree, the common value is separately continuous, and a separately continuous function is determined by its values on the dense set of actual states. Symmetry is the exchange of the two iterated limits. $\square$

Both readings. In the language of independent amalgams, existence is the existence of an independent amalgam, uniqueness is stationarity over a model (a type over a model is *stationary* if it has a unique nonforking extension to every larger set; in the iterated ultrapower, the type of the element realized second then does not depend on the ultrafilter), and symmetry is the symmetry of independence. For the energy comparison of two oscillators excited to infinite energy, the two orders of realization give the values 0 and 1, and the three properties are lost together.

The independent product is the double limit restated as a property of joint states: two idealized preparations combine in one way exactly when the two iterated limits agree.

Variatio 7. A 1 Ô VERO 2 CLAV. AL TEMPO DI GIGA

Discrimination and learning

Alice prepares one of many states, and Bob identifies it by adaptive probes, each chosen in view of the previous answers. In the online model of learning, a learner is shown a probe, predicts whether the response is high or low, and is then told the correct value. In this variation we bound the number of preparations that n probes can tell apart and the number of mistakes that the learner makes.

Trees. Let C be a self-adjoint bipartite observable, fix a margin $\delta > 0$, a threshold t and a sign $\theta = \pm 1$. A *tree of height n* assigns a probe σ_s to every node s of the complete binary tree of height n and a preparation ρ_η to every branch η , so that whenever η passes through s ,

$$\theta(\varphi_C(\rho_\eta, \sigma_s) - t) \geq \frac{\delta}{2} \text{ if } \eta \text{ turns right at } s, \quad \theta(\varphi_C(\rho_\eta, \sigma_s) - t) \leq -\frac{\delta}{2} \text{ if it turns left.}$$

A tree is an adaptive protocol: a sequence of n probes tells its 2^n preparations apart. A ladder with 2^h rungs contains a tree of height h : the first probe is the middle rung, and

binary search continues in the half of the ladder consistent with the outcome. The observable is *uniformly stable* if, at every margin, its ladders have bounded length.

Theorem 7.1 ([DI26a, Theorem B]). *For every bounded self-adjoint observable C , the following are equivalent:*

- (B1) *the observable C is uniformly stable;*
- (B2) (Shelah) *at every margin, the trees of C have bounded height;*
- (B3) *the expectation remains stable on ultrapowers, that is, on families of experiments;*
- (B4) (Sauer, Shelah) *at every margin δ , the number of preparations that n adaptive probes tell apart is at most a polynomial in n , whose degree is the bound on the height of trees;*
- (B5) (Littlestone) *at every margin, the response of an unknown preparation can be learned online with a bounded number of mistakes.*

If the conditions fail, then at some margin n probes tell apart 2^n preparations for every n .

Two arguments. The counting bound rests on an observation about the largest height $d(\mathcal{R})$ of a tree whose preparations lie in a set $\mathcal{R}$: at the first probe, $\mathcal{R}$ splits into the preparations that answer high and those that answer low, and the two parts cannot both carry trees of height $d(\mathcal{R})$: placed above the first probe, they would form a tree of height $d(\mathcal{R}) + 1$. By induction on n , with Pascal's rule, the number of preparations that n probes tell apart is at most $\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{d}$, a polynomial of degree d . This is the tree form of the lemma of Sauer and Shelah [Sau72, She72].

The learner uses the same observation [Lit88]. The learner maintains the set V of preparations consistent with the answers so far and predicts the side whose consistent preparations carry the higher tree. Each mistake decreases $d(V)$, so there are at most d mistakes. Conversely, along a tree of height n , choosing at each node the branch opposite to the prediction produces n mistakes.

For (B3), fix a nonprincipal ultrafilter $\mathcal{U}$ and consider sequences $x = (\rho_k)$ and $y = (\sigma_k)$, families of experiments indexed by k , with the expectation $\lim_{k, \mathcal{U}} \varphi_C(\rho_k, \sigma_k)$. Trees of every height glue along the ultrafilter into an infinite tree, whose $2^{\aleph_0}$ branches are separated by countably many probes. A stable expectation is a pairing with a reflexive space, and the closed span of countably many probes would then have a separable dual containing $2^{\aleph_0}$ separated points, which is impossible.

Example 7.2 (Stable but not uniformly stable). Let both systems carry the Hilbert space $\bigoplus_{n \geq 1} \mathbb{C}^n$, a sum of spin multiplets (only the orthonormal bases of the summands, the *sectors*, are used), with orthonormal bases $e_1^n, \dots, e_n^n$ of the summands, and let C_{sec} be 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 basis vectors form a ladder of n rungs at margin 1, so the observable C_{sec} is not uniformly stable. Nevertheless, it is stable: its effective observables factor through the ℓ^2 -sum of the spaces ℓ_n^1 , which is reflexive [DI26a, Example 9.8].

For the operator theorist. A tree of height n is a binary tree of Shelah for the formula φ_C , and the largest height of such trees is Shelah's binary tree rank; as observed by Chase and Freitag [CF19], it is the *Littlestone dimension* of the class of responses, the optimal mistake bound of (B5). By the theorem of Łoś, trees of unbounded height in $\mathcal{M}_C$ glue into an infinite tree in $\mathcal{M}_C^{\mathcal{U}}$, which is the equivalence of (B2) and (B3); hence uniform stability is stability

of φ_C in the theory of $\mathcal{M}_C$ (see Variatio 1), a property shared by all elementary extensions, and in Example 7.2 we show that it is stronger than stability of C . The count in (B4) is a count of φ -types over finite sets of probes, and Rowbottom's original definition of stability, at most λ types over λ parameters, is its infinite form; here a type over a set A is the type of an element of some elementary extension, and for countable A all of them are realized in $\mathcal{M}_C^u$.

For the logician. In the experiment behind Theorem 7.1, Alice supplies independent copies of an unknown preparation ρ chosen from a known list, and Bob, at each step, prepares a probe σ , measures C on $\rho \otimes \sigma$ repeatedly, and records whether the empirical mean lies above $t + \delta/2$ or below $t - \delta/2$; by the law of large numbers, the margin δ makes this decision reliable after finitely many repetitions. The choice of each probe may depend on the previous decisions, and a tree of height n is a protocol of this kind that identifies one of 2^n preparations with n decisions. For a uniformly stable observable with bound d on the height of trees, no protocol does this for $n > d$, and a learner who predicts each decision before it is made errs at most d times. In Example 7.2 the summand $\mathbb{C}^n$ is the state space of a particle of spin $(n-1)/2$, and uniformity fails because ever larger multiplets carry ever longer ladders.

The ultrapower expectation is a joint limit, not an iterated one, and uniform stability is the double limit criterion of Grothendieck applied to families of experiments: it holds exactly when trees have bounded height, and then adaptive discrimination grows polynomially and online learning makes boundedly many mistakes.

Variatio 8. A 2 CLAV.

Bounds independent of the dimension

The swap test and the coincidence of Hong, Ou and Mandel (see Variatio 4) are stable in every dimension. In this variation we show that they are uniformly stable, with bounds that do not depend on the dimension. Suppose that Bob prepares probe photons in arbitrary states and measures them adaptively against Alice's photon, choosing each probe in view of the previous outcomes. Combining Theorem 8.1 below with the counting bound of Variatio 7, we show that the number of Alice's preparations that n probes can tell apart is bounded by a polynomial in n whose degree depends only on the margin, not on the dimension: a photon with two polarizations, the internal space $K = \mathbb{C}^2$ of Variatio 4, and a photon with infinitely many spatial modes obey the same bound.

Decision trees. Bob's adaptive protocol is a binary decision tree. Fix a margin $\delta > 0$ and a bipartite expectation $\varphi(\rho, \sigma)$. A *tree of height n at margin δ* assigns a probe σ_s and a threshold t_s to every node s of the complete binary tree of height n , and a preparation ρ_b of Alice to every branch b , so that whenever the branch b passes through the node s ,

$$\varphi(\rho_b, \sigma_s) \geq t_s + \frac{\delta}{2} \text{ if } b \text{ turns right at } s, \quad \varphi(\rho_b, \sigma_s) \leq t_s - \frac{\delta}{2} \text{ if } b \text{ turns left.}$$

The 2^n preparations at the branches are then told apart by n probes, each answered with margin δ . Since both answers occur at every node, $|t_s|$ never exceeds the largest value of $|\varphi|$.

Theorem 8.1 (Novikoff [Nov62]; [DI26a, Theorem 9.11]). *Suppose that $\varphi(\rho, \sigma) = \langle u(\rho), v(\sigma) \rangle$ for maps u and v into a real Hilbert space $\mathcal{K}$ with $\|u(\rho)\| \leq R_u$ and $\|v(\sigma)\| \leq R_v$. Then every tree at margin δ has height at most $16(R_u R_v / \delta)^2$. For the swap test and for Hong–Ou–Mandel coincidences, with the margin measured in probability, the height is at most $4/\delta^2$.*

Proof. Write $x_s = (v(\sigma_s), -t_s/R_u)$ in the Hilbert space $\mathcal{K} \oplus \mathbb{R}$, so that $\|x_s\|^2 \leq 2R_v^2$. Walk down the tree from the root, keeping a vector w , initially 0. At the node s , turn left if

$\langle w, x_s \rangle \geq 0$ and right otherwise; set $y_s = +1$ for a right turn and $y_s = -1$ for a left one, and replace w by $w + y_s x_s$. Since $y_s \langle w, x_s \rangle \leq 0$ at every step, after n steps we have $\|w\|^2 \leq 2nR_v^2$. Let b be the branch that was followed and put $w^* = (u(\rho_b), R_u)$, so that $\|w^*\|^2 \leq 2R_u^2$ and $\langle w^*, x_s \rangle = \varphi(\rho_b, \sigma_s) - t_s$. By the definition of a tree, we have $y_s \langle w^*, x_s \rangle \geq \delta/2$ at every node of b , hence we have $\langle w, w^* \rangle \geq n\delta/2$. By the Cauchy–Schwarz inequality, we have $n\delta/2 \leq \sqrt{2n} R_v \cdot \sqrt{2} R_u$, that is, $n \leq 16(R_u R_v / \delta)^2$. For the swap test, the overlap $\text{Tr } \rho \sigma$ is the Hilbert–Schmidt inner product of two operators of norm at most 1, and a margin δ in probability is a margin 2δ in $\text{Tr } \rho \sigma$; the bound becomes $16/(2\delta)^2 = 4/\delta^2$. The same holds for coincidences. $\square$

At margin $\delta = 0.1$, no protocol of 401 beam-splitter probes tells apart 2^{401} photon states, however many modes the photons have. Let us notice that no such bound holds for arbitrary observables: for the comparison of two energies with discrete spectra, Bob’s probe f_l answers “yes” exactly on Alice’s states e_k with $k < l$, so binary search on the levels builds trees of every height at margin 1.

By the binary search of Variatio 7, a ladder of swap tests at margin δ has fewer than $2^{1+4/\delta^2}$ rungs.

For the operator theorist. The height of trees at margin δ is Shelah’s binary tree rank of the expectation, read as a formula, and Theorem 8.1 bounds it for every expectation that factors through a Hilbert space, with a bound that involves only the norms and the margin. By Theorem 7.1, such a bound is uniform stability, that is, stability of the expectation in the theory of $\mathcal{M}_C$ (see Variatio 1), and it persists in every elementary extension, in particular in every ultrapower. The proof is the perceptron algorithm, an online learner in the sense of (B5), run along the branch chosen against its predictions, and the bound on its number of mistakes is the bound on the rank.

For the logician. The internal space K of a photon is $\mathbb{C}^2$ when only polarization is considered and infinite-dimensional once spatial or temporal modes are included, as they are for any photon produced in a laboratory. A probe of Theorem 8.1 is the experiment of Variatio 4, repeated on copies: Bob sends a photon in a state σ_s of his choice into the second port of the beam splitter and estimates the probability $\frac{1}{2}(1 - \text{Tr } \rho \sigma_s)$ that the two photons leave by different ports. The theorem asserts that, at margin δ in this probability, no adaptive sequence of more than $4/\delta^2$ such experiments tells apart all the preparations of a tree, whatever the dimension of K . In each fixed finite dimension some bound follows from the compactness of the state space; the content of the theorem is that the bound does not grow with the dimension.

If the two iterated limits $\lim_n \lim_m \varphi(\rho_n, \sigma_m)$ and $\lim_m \lim_n \varphi(\rho_n, \sigma_m)$ of an overlap measurement differed, subsequences would give ladders with any number of rungs at a fixed margin. Thus, for overlap measurements the two orders of idealization agree, and the theorem turns this agreement into a count: at margin δ , the staircase of the introduction cannot be built from overlaps with more than $2^{1+4/\delta^2}$ steps.

Variatio 9. A 1 CLAV. CANONE ALLA TERZA

Mixtures simulate idealized preparations

An oscillator excited to infinite energy cannot be prepared in a laboratory, but a mixture $\sum_k t_k \rho_k$ of actual preparations, which produces ρ_k with probability t_k , can. In this variation we show that, for stable observables, finite mixtures of actual preparations simulate idealized

ones uniformly over all probes. The logician averages formulas and the physicist averages preparations; since the expectation is affine in the state, the two averages coincide.

The logician. In real-valued logic the definition of a stable φ -type over a model M is a uniform limit of continuous combinations of instances $\varphi(a, y)$ with a in M [BYU10]. Ben Yaacov observed that, by the theorem of Mazur, averages suffice: the definition is a uniform limit of $\frac{1}{n} \sum_{i \leq n} \varphi(a_i, y)$ [BY14].

The physicist. Let C be a bipartite observable, let ω_1 be an idealized state of Alice, and let $\otimes_B$ denote the order of idealization with Bob's system idealized first.

Theorem 9.1 ([DI26a, Theorems 6.1 and 6.5]).

- (a) *The observable C is stable if and only if, for every idealized state ω_1 of Alice and every $\varepsilon > 0$, some density operator ρ satisfies $|(\omega_1 \otimes_B \omega_2)(C) - \omega_2(T_C \rho)| \leq \varepsilon$ for all states ω_2 of Bob, idealized or not. The density operator ρ can be taken to be a finite mixture from the tail of any sequence converging to ω_1 .*
- (b) *Suppose that $\varphi_C(\rho, \sigma) = \langle a(\rho), b(\sigma) \rangle$ for maps into a Hilbert space with $\|a(\rho)\| \leq R_a$, where b extends to a bounded map of norm $\|b\|$ on trace-class operators. Then equal-weight mixtures of any $N \geq (2R_a\|b\|/\varepsilon)^2$ suitably chosen tail preparations suffice. For the swap test and the coincidence of Hong, Ou and Mandel, $N \geq 1/\varepsilon^2$ suffices, in every dimension.*

Idea of the proof. (a) If C is stable, the response of ω_1 is an observable Z , and it lies in the weak closure of the effective observables of mixtures of any tail. By the theorem of Mazur [Maz33], the weak and norm closures of a convex set coincide, so some mixture ρ has $\|Z - T_C(\rho)\| \leq \varepsilon$, and a norm approximation is uniform over all states ω_2 . Conversely, if the approximation holds for every ε , the idealized response of ω_1 is a norm limit of observables, hence an observable, and C is stable because every idealized response is an observable.

(b) This is Maurey's empirical method [Pis81]. Let $\xi = \sum_k t_k a(\rho_k)$ approximate the weak limit of $a(\rho_k)$ within $\varepsilon/(2\|b\|)$, as in (a). Draw N independent indices with probabilities t_k . The average $\bar{\xi}$ of the drawn vectors satisfies $\mathbb{E}\|\bar{\xi} - \xi\|^2 = \frac{1}{N} \mathbb{E}\|a(\rho_k) - \xi\|^2 \leq R_a^2/N$, so some draw has $\|\bar{\xi} - \xi\| \leq R_a/\sqrt{N} \leq \varepsilon/(2\|b\|)$. For the swap, a and b are the identity into the Hilbert-Schmidt operators, with $R_a = \|b\| = 1$, which gives $N \geq 4/\varepsilon^2$; the swap test accepts with probability $\frac{1}{2}(1 + \text{Tr}(\rho\sigma))$, so its errors are half those of the swap, and $N \geq 1/\varepsilon^2$ suffices. $\square$

The logician. The order of idealization in (a) is the order in which the definition is evaluated: Bob's type is substituted into the definition of Alice's.

The physicist. Let us notice that the order cannot be dropped: let $(e_k)_{k \geq 0}$ and $(f_l)_{l \geq 1}$ be orthonormal bases, let S_k project onto the span of $f_1, \dots, f_k$, and let $C' = |e_0\rangle\langle e_0| \otimes I + \sum_{k \geq 1} |e_k\rangle\langle e_k| \otimes S_k$. Then the value $\varphi_{C'}(e_k, f_l)$ is 1 exactly when $l \leq k$, so C' is unstable. Nevertheless, for $\omega_1 = \lim_{k, \mathcal{U}} e_k$ the actual state e_0 reproduces ω_1 exactly with Alice idealized first, since $T_{C'}(e_0) = I$ is the response of ω_1 on every actual probe. With Bob idealized first, at $\omega_2 = \lim_{l, \mathcal{U}} f_l$, the idealized oscillator gives 0 and e_0 gives 1. For the plain energy comparison the failure is total: no mixture comes within any error less than 1, even on actual probes [DI26a, Example 6.3].

Both readings. In general, the definitions of real-valued logic are polynomials or lattice combinations of instances. For an observable an average of instances is again an instance, $\sum_k t_k \varphi_C(\rho_k, \cdot) = \varphi_C(\sum_k t_k \rho_k, \cdot)$, so the approximating formula is a physical preparation, and polynomials of degree one suffice.

The double limit is what places the idealized response in the weak closure of actual responses as an observable; Mazur's theorem then turns the weak limit into a norm limit, so that every idealized preparation is approximated, uniformly over all probes, by a finite mixture that a laboratory can prepare.

Variatio 10. A 1 CLAV. FUGHETTA

Countable randomization does not detect instability

Although an oscillator of infinite energy cannot be prepared, its level can be drawn at random. Alice draws her level k from a probability distribution μ , Bob draws his level l from a distribution ν , and the energy comparison is averaged. If Alice's average is taken first, the result is $\int(\int \varphi(k, l) d\mu(k)) d\nu(l)$; if Bob's is taken first, the integrals are exchanged. In this variation we show that the two iterated averages agree for every observable when the distributions are countably additive, and that they agree for all finitely additive means exactly when the observable is stable.

The question is whether the two iterated integrals agree; we consider it for three classes of measures. It is convenient to integrate over all states: a *mean* on $\mathfrak{D}(H)$ is a finitely additive probability measure defined on all of its subsets, and its *barycenter* is the state $A \mapsto \int \text{Tr}(\rho A) d\mu(\rho)$. Since the expectation of a bipartite observable C is affine in each state, an iterated integral of φ_C depends only on the barycenters, and it is one of the two products of the barycenters:

$$\int\left(\int \varphi_C d\mu\right) d\nu = (\omega_\mu \otimes_{\mathbf{A}} \omega_\nu)(C), \quad \int\left(\int \varphi_C d\nu\right) d\mu = (\omega_\mu \otimes_{\mathbf{B}} \omega_\nu)(C).$$

Finite mixtures. If μ is finitely supported, its barycenter is a mixture, a density operator, and finite sums commute. The two integrals agree for every observable.

Countably additive measures. Let μ be a countably additive Borel probability measure on $\mathfrak{D}(H_1)$, with the trace norm. Its barycenter is the Bochner integral $\int \rho d\mu(\rho)$, again a density operator. The two products agree whenever one factor is actual, so the two iterated integrals agree for every observable and every mean ν of Bob [DI26a, Proposition 7.2]. No countably additive randomization detects instability, for any observable.

Finitely additive means. For finitely additive means, in Theorem 10.1 below we show that the answer is stability.

Theorem 10.1 (Pták, Simons; [DI26a, Theorem 7.1]). *An observable C is stable if and only if, for all means μ on $\mathfrak{D}(H_1)$ and ν on $\mathfrak{D}(H_2)$,*

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

Proof. The two sides are the two products of the barycenters. Every state is a barycenter: if ω is the limit of a net (ρ_i) of density operators, and $\mathcal{W}$ is an ultrafilter on the index set containing all tails, then the function μ given by $\mu(g) = \lim_{i, \mathcal{W}} g(\rho_i)$ is a mean with barycenter ω . Hence the identity holds for all means exactly when the two orders of idealization agree for all states, which is stability. $\square$

For the energy comparison, let μ and ν be “uniform distributions on all levels”: invariant means on $\mathbb{N}$, such as Banach limits of the uniform distributions on $\{1, \dots, N\}$, placed on Alice's states e_k and on Bob's states f_l . For each fixed l , almost every k in the sense of μ exceeds l , so the inner integral over Alice's levels is 0; for each fixed k , almost every l in the sense of ν exceeds k , so the inner integral over Bob's levels is 1. The two iterated integrals are 0 and 1. The same happens for ultrafilters, the $\{0, 1\}$ -valued means: for a nonprincipal

ultrafilter $\mathcal{W}$, the set $\{(k, l) : k < l\}$ has measure 0 under $\lim_{l, \mathcal{W}} \lim_{k, \mathcal{W}}$ and measure 1 under $\lim_{k, \mathcal{W}} \lim_{l, \mathcal{W}}$. No countably additive version exists, because a countably additive distribution on $\mathbb{N}$ cannot give every finite set measure zero.

For the operator theorist. A mean μ on Alice's state space defines a *Keisler measure* over $\mathcal{M}_C$ (see Variatio 5), the functional $\theta(x, b) \mapsto \int \theta(\rho, b) d\mu(\rho)$ on formulas with parameters from $\mathcal{M}_C$; a mean that takes only the values 0 and 1 is an ultrafilter, and its Keisler measure is the type of an element of an ultrapower. The two iterated integrals of Theorem 10.1 are the two products of Keisler measures, the averaged form of the two independent amalgams of Variatio 1, and the theorem says that these products commute for all measures exactly when they commute for all types. A countably additive measure on $\mathfrak{D}(H_1)$ with the trace norm has a density operator as barycenter, an element of $\mathcal{M}_C$ itself; for the formula φ_C it behaves like a realized type, and the two products agree because they agree whenever one factor is realized.

For the logician. A laboratory that draws Alice's preparation at random, by a countably additive distribution μ , and does not record the draw, prepares the density operator $\int \rho d\mu(\rho)$, since the expectation of every observable is affine in the state. Hence, if either system is prepared by a countably additive randomization, that system is in an actual state, and the two iterated averages agree for every observable, because the two orders of idealization agree when one factor is actual (see Variatio 1). The disagreement found for the energy comparison needs Alice's level to be distributed uniformly over all of $\mathbb{N}$, a finitely additive mean that vanishes on every finite set of levels; its barycenter is the oscillator excited to infinite energy, an idealized state that no random procedure produces.

The double limit of Grothendieck is thus a Fubini theorem. For finite mixtures and countably additive measures it holds for every observable; for finitely additive means it holds exactly for the stable ones.

Part II. NIP

Variatio 11. A 2 CLAV.

The independence property

A photon of wavenumber k that passes a double slit is detected at position x on Bob's screen with probability density proportional to $1 + \cos kx$, in suitable units, so that the fringes have spacing $2\pi/k$. If photons are prepared with wavenumbers $8, 8^2, 8^3, \dots$, then, as we show in Variatio 16, a single position of Bob's detector can be bright for any chosen finite set of them and dark for any other finite set disjoint from it: the fringes realize every pattern. The energy comparison cannot do this, because its answers are monotone. In this variation we draw the second dividing line, between the instability of the energy comparison, which is NIP, and instability with the independence property, and name the configuration that it forbids.

A family of real functions on a set Y has the *independence property* if there are a sequence (r_n) in it and reals $a < b$ such that every finite pattern is realized: for all finite disjoint $E, F \subseteq \mathbb{N}$, some $y \in Y$ has $r_n(y) \leq a$ for $n \in E$ and $r_n(y) \geq b$ for $n \in F$. An observable is *NIP* if Alice's responses do not have it. In Theorem 11.1 below, we show that stable observables are NIP; the energy comparison is NIP, since its answers are monotone, but unstable. Unlike the order property, the independence property depends on which system is Alice's: a spin chain whose sites are probed one at a time (see Variatio 14) is NIP for Alice's

responses but not for Bob's [DI26b, Corollary 8.7]. In model theory a formula is NIP exactly when its opposite, with the roles of the variables exchanged, is NIP [She90]. For observables the uniform version, which bounds the size of finite independent sets, is still symmetric, but NIP itself, which forbids only an infinite independent sequence, is not.

Theorem 11.1 ([DI26b, Theorem 6.1]). *Let C be a bipartite observable.*

- (a) *The observable C is stable if and only if every idealized response is continuous on the space of probes.*
- (b) *The observable C is NIP if and only if every idealized response is of the first Baire class. The idealized responses then form a Rosenthal compactum.*
- (c) *If C is not NIP, then the idealized responses contain a homeomorphic copy of $\beta\mathbb{N}$.*

The theorem is the dichotomy of Rosenthal and of Bourgain, Fremlin and Talagrand [Ros74, BFT78], applied to the space of probes, the state space of the C^* -algebra generated by the effective observables.

For the operator theorist. For a formula in a theory, the independence property says that for every n some n parameters admit all 2^n patterns; by compactness, an infinite independent sequence then exists in an elementary extension, such as an ultrapower, whereas for the single structure $\mathcal{M}_C$ the definition above asks for an infinite independent sequence in $\mathcal{M}_C$ itself. Theorem 11.1 describes the φ -types over $\mathcal{M}_C$ realized in ultrapowers, the idealized responses of Variatio 1: for an NIP observable each is the pointwise limit of a sequence of responses of actual preparations, as Simon proved for NIP formulas [Sim15], while an independent sequence (ρ_n) yields the copy of $\beta\mathbb{N}$, formed by the φ -types of the elements $[\rho_n]_{\mathcal{V}}$ for the ultrafilters $\mathcal{V}$ on $\mathbb{N}$. Because the condition concerns infinite sequences in $\mathcal{M}_C$ and not finite configurations in every model of its theory, it need not be symmetric in Alice and Bob, whereas NIP of a formula in a theory, like the uniform version for observables, is symmetric.

For the logician. In the double-slit experiment Bob measures the position at which the photon is detected, and the response of a preparation of wavenumber k to a detector at x is the relative intensity $\frac{1}{2}(1 + \cos kx)$ of the fringes there. The independence property of the fringes is a statement about single detector positions: for the wavenumbers $8, 8^2, 8^3, \dots$, the intensities at x are close to 1 or close to 0 according to the base-8 digits of $x/2\pi$, which can be chosen freely (see Variatio 16). For the comparison of photon numbers, by contrast, Bob's state f_l answers yes exactly for Alice's levels $k < l$, so on any n levels only the $n + 1$ initial segments occur as patterns. In scanning a spin chain the roles are reversed: the chains of Alice realize every pattern on Bob's sites, which is why NIP for Alice's responses does not give NIP for Bob's.

When the double limit fails, there are two cases. If the observable is NIP, every idealized response is the limit of a sequence of actual responses; otherwise the idealized responses contain a copy of $\beta\mathbb{N}$, whose nonprincipal points are not limits of sequences of actual responses.

Variatio 12. CANONE ALLA QUARTA

The rational order

The energy comparison of Part I compares two discrete quantities; its continuous counterpart, the comparison of two positions on a line, is governed by the rational order, in which every finite linear order embeds. In this variation we approach the rational order from two

directions: the logician descends from the theory of the rational order to its types, and the physicist ascends from two particles on a line to the idealized limits of their comparison.

The finite linear orders form a Fraïssé class, and their limit is $(\mathbb{Q}, <)$, the dense linear order without endpoints. Its theory is NIP and unstable, the prototype of a theory that is ordered but tame. The types of the formula $x < y$ over the real line are the cuts, and each real number r splits into two of them, just below and just above r ; in $(\mathbb{R}, <)^u$ the class of $(r + 1/n)$ realizes the type just above r and the class of $(r - 1/n)$ the type just below. The space of these types is the split interval of Alexandroff, which is hereditarily separable but not metrizable. The physical home of the rational order is the comparison of two continuous quantities, such as the positions of two particles on a line.

Over a finite set of n parameters the formula $x < y$ has only $n + 1$ φ -types, since a new element falls into one of the $n + 1$ gaps that the parameters leave. A set of n parameters is *shattered* by φ if each of its subsets is cut out by $\varphi(a, y)$ for some a , and the *Vapnik–Chervonenkis dimension* of φ is the largest n for which some n parameters are shattered; by the theorem of Łoś, unbounded shattering in M gives the independence property in M^u . This is NIP in its finite form: the number of types grows polynomially, as the lemma of Sauer and Shelah demands of a formula of Vapnik–Chervonenkis dimension one, whereas a formula with the independence property has 2^n types over some set of n parameters. Over an infinite set the count changes character. The types are the cuts of the set, and their number is governed by the cardinal $\text{ded } \kappa$, the supremum of the numbers of cuts of linear orders of size κ ; Shelah showed that over a set of size κ , with $|T|$ the cardinality of the language plus $\aleph_0$, a stable theory has at most $\kappa^{|T|}$ complete types and an NIP theory at most $(\text{ded } \kappa)^{|T|}$ [She90, CS16]. The passage from stability to NIP is the passage from points to cuts.

Independence of types changes accordingly. In an NIP theory, a global type does not fork over a model M exactly when it is invariant over M , that is, fixed by every automorphism of the monster model that fixes M pointwise [HP11]; the coheirs of Variatio 1 are of this kind. Unlike the stable case of Variatio 6, a type over a model may have several invariant extensions. Let us notice that the rational order shows two of them for the type just below a real number r : realized in an iterated ultrapower, the type just below r for Alice and for Bob gives two independent amalgams: if Alice's particle, spread over $(r - 1/n, r)$, is idealized first, it lies closer to r than Bob's, spread over $(r - 1/m, r)$, and the comparison “Alice's particle lies to the left of Bob's” takes the value 0; in the other order it takes the value 1. This is the staircase of the introduction, compressed into a single point.

For the operator theorist. For the formula $x < y$, the ultrapower $(\mathbb{R}, <)^u$ exhibits the types: over $\mathbb{R}$, the element $[r + 1/n]$ has the φ -type $y \mapsto 1_{(r, \infty)}(y)$ and the element $[r - 1/n]$ the φ -type $y \mapsto 1_{[r, \infty)}(y)$, two functions of the parameter that differ only at $y = r$. Together with the two types of infinite elements, these are all the φ -types over $\mathbb{R}$, and each is a pointwise limit of a sequence of realized types, as Theorem 11.1(b) requires of an NIP formula. Since the two types of such a pair are separated only by the parameter r itself, no countable set of parameters separates all of them, and the type space, the split interval, is hereditarily separable but not metrizable.

For the logician. For two particles on a line the comparison is the relation observable of $\{(x, y) : x < y\}$, multiplication by its indicator on $L^2(\mathbb{R}^2)$, and in states ρ and σ it answers yes with probability $(\mu_\rho \otimes \mu_\sigma)(\{x < y\})$, where μ_ρ and μ_σ are the distributions of the two positions. If Alice's particle is spread over $(r, r + 1/n)$ and n grows, its responses converge to $\sigma \mapsto \mu_\sigma((r, \infty))$, and if it is spread over $(r - 1/n, r)$, to $\sigma \mapsto \mu_\sigma([r, \infty))$. On actual probes the

two limits agree, since a density operator gives probability zero to the point r ; they disagree on the idealized probe obtained by spreading Bob's particle over $(r - 1/m, r + 1/m)$ and letting m grow first, where the left limit answers 1 and the right limit answers 0. These two idealizations of a particle at r form a split pair.

The rational order thus fails the double limit in the mildest way: at each point of the line, the limits of a response from the left and from the right may disagree, these one-sided limits form the split pairs, and every type remains of the first Baire class.

Variatio 13. A 2 CLAV.

The dictionary of NIP

For an NIP formula φ and a countable model M , Simon showed that the φ -types over a monster model that are finitely satisfiable in M , viewed as functions on the space of types of parameters over M , are of the first Baire class, each the limit of a sequence of types realized in M [Sim15]. In the language of Variatio 1, such a type is $b \mapsto \lim_{n,\mathcal{U}} \varphi(\rho_n, b)$ for b in any elementary extension: Alice's idealized response, evaluated with Bob idealized first. Bipartite observables present the same situation, with Bob's probes as parameters and Alice's idealized responses as types; these types are coheirs, hence, for an NIP observable, nonforking extensions in the sense of invariance of Variatio 12. An NIP observable is NIP_3 , NIP_2 or NIP_1 when its idealized responses form a metrizable, a hereditarily separable or a first countable compactum. In this variation we set out the second dictionary (see Table 3) and describe each of these levels by how many probes, and which, are needed to separate the idealized responses.

Two facts about Rosenthal compacta carry the dictionary. The first is the theorem of Bourgain, Fremlin and Talagrand that a Rosenthal compactum is *Fréchet–Urysohn*: every point in the closure of a set is the limit of a sequence from the set [BFT78]. For an NIP observable, an idealized response $\lim_{n,\mathcal{U}} r_{\rho_n}$, where r_ρ denotes the response $\sigma \mapsto \varphi_C(\rho, \sigma)$ of the preparation ρ , is therefore also $\lim_k r_{\rho_{n_k}}$ for a subsequence, so that the ultrafilter, indispensable in Variatio 1, can be replaced by a sequence. For the fringes this fails: the closure of their responses contains $\beta\mathbb{N}$, which has no nontrivial convergent sequences, and the points of $\beta\mathbb{N} \setminus \mathbb{N}$ are ultrafilter limits that no sequence reaches. The second fact is that the closure is determined by the responses: its topology is pointwise convergence on Bob's probes, so a level of the trichotomy is a statement about how many probes, and which, decide convergence.

A single example shows both facts at work. Let Alice's particle on a line be spread over $(r, r + 1/n)$. Its response to a probe σ is the probability $(\mu_n \otimes \mu_\sigma)(\{x < y\})$, and as n grows it converges, at every probe, to $\sigma \mapsto \mu_\sigma((r, \infty))$: an ultrafilter is not needed. In the ultrapower, the class of these preparations has as its φ -type the cut just above r ; the class of the particles spread over $(r - 1/n, r)$ has the cut just below. On actual probes the two types agree, because a density operator gives probability zero to the point r ; they are told apart only by idealized probes concentrated at r , with Bob idealized first. The responses are monotone in the position of Bob's particle, each with at most countably many jumps, and they are functions of the first Baire class, as Simon's theorem requires.

Metrizable compacta (NIP_3). A compact set of functions is metrizable exactly when countably many points separate it [DI26b, Proposition 4.5]. For types, this says that one fixed countable set of instances of the formula decides every limit type. For an observable, it says that countably many probes tell every idealized response apart. A mode of light is an oscillator whose eigenvector e_n is the state with n photons, so the comparison of photon numbers

Model theory and topology	Quantum measurement
the independence property	a sequence of preparations from which probes extract every pattern: interference fringes
NIP	comparisons of physical quantities, and the causal order
Baire class one (Rosenthal; Bourgain, Fremlin, Talagrand)	every idealized response is the limit of a sequence of actual responses
a Rosenthal compactum metrizable, and for observables hereditarily separable	the idealized responses of an NIP observable countably many probes separate the idealized responses
first countable, not hereditarily separable	a jump approached from either side
a split pair	the two pure phases of a ferromagnet
the Helly space	the idealized responses of a comparison of positions
$\hat{A}(2^{\mathbb{N}})$, not first countable	a light cone shrinks to a point
the seven minimal families	trees of preparations of photons, of a particle in a box, of a spin chain
symmetry of NIP in the two variables	holds for uniform NIP; fails for NIP: a spin chain read site by site

TABLE 3. The dictionary of NIP.

is the energy comparison of the staircase. This comparison is of this kind, NIP_3 , because its responses can jump only at the countably many levels of the spectrum.

Hereditarily separable compacta (NIP_2). A separable Rosenthal compactum is hereditarily separable exactly when it has no discrete subset of the size of the continuum [Tod99]. For types, there is no continuum of limit types each of which finitely many instances tell apart from all the others. The comparison of positions fails this condition: the midpoints of its split pairs, one for each point of the line, form a discrete set: each is isolated by the probe at its own point, where it alone takes the value $\frac{1}{2}$. For observables the condition and metrizability coincide (see Variatio 19), which is why NIP_2 and NIP_3 merge.

First countable compacta (NIP_1). A point of a compact space of functions has a countable base of neighborhoods exactly when, for some countable set of coordinates, it is the only point of the compactum that agrees with it there [DI26b, Proposition 4.6]. For types, each limit type is decided by countably many instances of its own, chosen for it. The comparison of positions is of this kind, NIP_1 : a response that jumps at a point is decided by the probes at that point and at rational points around it.

Beyond first countability. The causal order and the comparison of phases have limit types that no countable set of instances decides: a light cone that has shrunk to a point c is distinguished from nearby responses only by probes that give positive mass to c , and a countable set of probes gives positive mass to only countably many points.

Read in this way, the classes measure how far an idealized response may be from continuous while remaining a limit of actual ones. Continuity, the condition of Grothendieck’s double limit, is the strongest of them: a continuous response is determined at each probe by the probes near it.

Variatio 14. A 2 CLAV.

The Todorčević trichotomy

When countably many probes no longer separate the idealized responses, the failure takes one of three specific forms [Tod99]. In this variation we state the trichotomy of Todorčević for separable Rosenthal compacta and list, in Table 4, a physical representative of each class. The table anticipates two later results: in Theorem 16.1 of Variatio 16 we place the representatives in their classes, and in Theorem 19.1 of Variatio 19 we show that the hereditarily separable, non-metrizable class is empty for observables.

Recall that the idealized responses are metrizable exactly when countably many probes separate them [DI26b, Proposition 4.5]. A separable Rosenthal compactum that is not metrizable contains a homeomorphic copy of one of a few canonical compacta: one of the split Cantor sets $\hat{S}_{\pm}(2^{\mathbb{N}})$ if it is hereditarily separable; one of the duplicates $\hat{D}(2^{\mathbb{N}})$ and $\hat{D}(S(2^{\mathbb{N}}))$ if it is first countable but not hereditarily separable; and the compactum $\hat{A}(2^{\mathbb{N}})$, built over the one-point compactification of a discrete set of cardinality $2^{\aleph_0}$, if it is not first countable. Figure 4 draws the trichotomy as a decision tree.

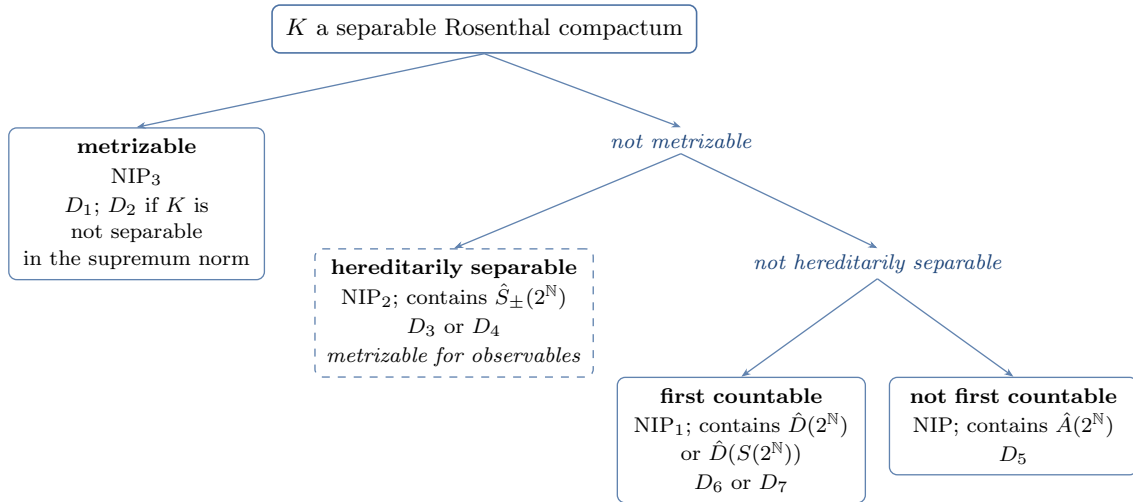


FIGURE 4. The Todorčević trichotomy as a decision tree, with the canonical compacta and the minimal families of the heptachotomy at its leaves (compare [DIMW⁺26, Section 4.2]). For observables the dashed leaf merges with the metrizable one (see Theorem 19.1).

The observables of Part II. For the logician, we define once the observables that represent the classes, and the other observables that Part II classifies. A particle on a measure space X is described by $L^2(X)$, and its position is the family of multiplication operators M_{1_E} : a state ρ places the particle in a Borel set E with probability $\mu_\rho(E) = \text{Tr}(\rho M_{1_E})$. For particles of Alice and of Bob on X and on Y and a Borel set $R \subseteq X \times Y$, the *relation observable* C_R is multiplication by 1_R on $L^2(X) \otimes L^2(Y) = L^2(X \times Y)$, the event “Alice’s value and

Class	Witness	The picture	The physics
stable	continuous	idealizations commute	swap test
NIP_3	metrizable	countably many probes suffice	photon numbers; spin chains
$\text{NIP}_2 \setminus \text{NIP}_3$	split interval	mixtures supply the midpoints	empty for observables
$\text{NIP}_1 \setminus \text{NIP}_2$	duplicate	approach from either side	positions; quadratures; the Ising model
$\text{NIP} \setminus \text{NIP}_1$	$\hat{A}(2^{\mathbb{N}})$	a light cone shrinks to a point	the causal order; phases on a circle
not NIP	$\beta\mathbb{N}$	ever finer oscillation	interference fringes

TABLE 4. The classes of the Todorćević trichotomy and their physical representatives.

Bob's value are related by R'' , and its expectation is $\varphi_{C_R}(\rho, \sigma) = (\mu_\rho \otimes \mu_\sigma)(R)$ [DI26b, Lemma 7.1]. For countable sets A and B and a function $\varphi: A \times B \rightarrow [0, 1]$, the *diagonal observable* $C_\varphi = \sum_{a,b} \varphi(a, b) |a\rangle\langle a| \otimes |b\rangle\langle b|$ on $\ell^2(A) \otimes \ell^2(B)$ reads $\varphi(a, b)$ on the pair of labels (a, b) .

- *Photon numbers.* A mode of light is an oscillator, and its eigenvector e_n is the state with n photons; the comparison of photon numbers is the energy comparison P of the staircase, the relation observable of $\{k < l\}$ on $\mathbb{N} \times \mathbb{N}$.
- *Positions.* The comparison of positions is the relation observable of $\{x < y\}$ for particles on a line, or in a box $[0, 1]$; a particle *spread over* an interval I is in the state $|I|^{-1}|1_I\rangle\langle 1_I|$. The *quadratures* of a mode are $Q_\theta = \cos \theta Q - i \sin \theta d/dq$ on $L^2(\mathbb{R})$, with Q the position, each unitarily equivalent to Q , so comparing quadratures is unitarily equivalent to comparing positions.
- *Light cones.* In units in which light has speed 1, an event (t, ξ) of a $(1 + 1)$ -dimensional spacetime with $t \leq 0$ can influence a detection at the point y of Bob's time slice $t = 0$ exactly when $|y - \xi| \leq -t$, that is, when $a \leq y \leq b$ for the points $a = \xi + t$ and $b = \xi - t$ where the two light rays through the event cross the slice. The causal-order observable is the relation observable of $\{(a, b, y) : a \leq y \leq b\}$ on $[0, 1]^2 \times [0, 1]$.
- *Phases.* For particles on a circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, the phase comparison is the relation observable of $\{(x, y) : y - x \in (0, \frac{1}{2}) \bmod 1\}$, the event "Bob's angle is ahead of Alice's by less than half a turn".
- *Spin chains.* Alice holds a chain of spins of variable finite length, with Hilbert space $\ell^2(2^{<\mathbb{N}})$; the basis vector e_t is the chain of length $|t|$ whose spin at site n is up exactly when $t(n) = 1$. In *scanning*, or reading site by site, Bob's label is a site n and the observable reads 1 when Alice's spin at n is up. In *template matching*, Bob's label is a finite reference configuration u and the observable reads 1 when Alice's chain is an initial segment of u . The *combined* observable offers Bob both kinds of label. All three are diagonal observables.
- *Fringes.* The *wavenumber* of monochromatic light is $k = 2\pi/\lambda$, with λ the wavelength. Behind two slits, the two paths to a point x of a distant screen differ in phase by kx , in suitable units, so the amplitudes add to a multiple of $1 + e^{ikx}$ and the probability density is proportional to $|1 + e^{ikx}|^2 = 2(1 + \cos kx)$. Alice's register holds the wavenumber

and Bob's the detector position, and the fringe observable is $C = \sum_n |e_n\rangle\langle e_n| \otimes M_{f_n}$ on $\ell^2(\mathbb{N}) \otimes L^2[0, 2\pi]$, with $f_n(x) = \frac{1}{2}(1 + \cos k_n x)$ for a sequence (k_n) of wavenumbers.

- *The Ising family.* The example of Variatio 17 is a family of response functions, indexed by finite samples with offsets and by applied fields, to which the classification applies through the pointwise closure of the family.

For the operator theorist. The compactum classified here is the space of idealized responses, that is, of the φ -types over a large elementary extension of $\mathcal{M}_C$ that are finitely satisfiable in $\mathcal{M}_C$ (see Variatio 13), the coheirs of Variatio 1; by Theorem 11.1 it is a Rosenthal compactum when C is NIP. Its topology is that of pointwise convergence on the parameters, so metrizability says that the values at countably many parameters determine every such type, and first countability says that each type is determined, among all of them, by countably many parameters chosen for it. When metrizability fails, the trichotomy names the configuration of types that must be present: split pairs of types that differ at a single parameter, an uncountable discrete family of types, or a single type that is the limit of continuum many types isolated from one another.

For the logician. Metrizability has an experimental meaning: there is a fixed countable list of Bob's probes whose answers determine the response of every idealized preparation of Alice to every probe. For the comparison of photon numbers such a list exists (see Theorem 16.1), because a response can jump only at the countably many points of the spectrum. For the comparison of positions no countable list suffices, since an idealized response can jump at any point of the line, but each single idealized response is determined by probes at countably many points chosen for it. For the causal order even this fails: an event whose light cone has shrunk to the point c of Bob's time slice is distinguished only by probes that give positive probability to c itself (see Variatio 16).

Each class of the trichotomy is thus a degree of failure of the double limit that stops short of the independence property, and for observables one of the classes is empty (see Theorem 19.1).

Variatio 15. A 1 CLAV. CANONE ALLA QUINTA. ANDANTE

The heptachotomy

A tree of preparations branches in two at every node, and, along a dyadic subtree, the preparations on each branch converge to idealized limits. The logician classifies the trees that converge to a given compactum; the physicist identifies the compacta reached by trees of photons, particles and spins. The trichotomy says which canonical compactum appears; the finer classification of Argyros, Dodos and Kanellopoulos says how it appears. In this variation we state the heptachotomy, a refinement of the trichotomy along trees, and realize each of its seven minimal families by a tree of physical preparations; the observables used here are placed in their classes by Theorem 16.1 of the next variation.

The heptachotomy describes families of functions indexed by the dyadic tree $2^{<\mathbb{N}}$. Two families $\{f_t\}$ and $\{g_t\}$ indexed by $2^{<\mathbb{N}}$ are *equivalent* if the map $f_t \mapsto g_t$ extends to a homeomorphism of their pointwise closures, and a family that is relatively compact in the first Baire class is *minimal* if it is equivalent to its restriction to every dyadic subtree. There are exactly seven minimal families $D_1, \dots, D_7$ up to equivalence, and every family indexed by $2^{<\mathbb{N}}$ that is relatively compact in the first Baire class of a Polish space contains, along a regular dyadic subtree, a family equivalent to one of them [ADK08, Theorem 2]. Their closures fall into the

classes of the trichotomy: those of D_1 and D_2 are metrizable, those of D_3 and D_4 are the split Cantor sets, those of D_6 and D_7 are the duplicates, and that of D_5 is $\hat{A}(2^{\mathbb{N}})$.

Family	Closure	The picture	The physics
D_1	$A(2^{<\mathbb{N}})$	escape to infinity	photon number states
D_2	$2^{\leq \mathbb{N}}$	infinite configurations	a spin chain read site by site
D_3	$\hat{S}_+(2^{\mathbb{N}})$	approach from the right	a particle in a box
D_4	$\hat{S}_-(2^{\mathbb{N}})$	approach from the left	a particle in a box
D_6	$\hat{D}(2^{\mathbb{N}})$	twins over a Cantor set	a spin chain, sites and templates
D_7	$\hat{D}(S(2^{\mathbb{N}}))$	approach from both sides	a particle in a box, half on each side
D_5	$\hat{A}(2^{\mathbb{N}})$	a point-supported limit	light cones; templates

TABLE 5. The seven minimal families, grouped by class of the Todorćević trichotomy.

Theorem 15.1 ([DI26b, Theorem 8.5]). *Each of the seven minimal families is realized by a tree of pure preparations of a physically natural observable: the family D_1 by photon number states, compared by photon number; the family D_2 by a spin chain read site by site; the families D_3 and D_4 by a particle in a box spread over a gap of a Cantor set, compared by position; the family D_5 by a spin chain matched against templates, and also by the causal order of spacetime; the family D_6 by a spin chain read both site by site and against templates; and the family D_7 by a particle in a box spread half over a short interval on each side of a point.*

The realizations display what distinguishes the families (see Figure 5). The responses of photon number states are isolated points with a single idealized limit, their only accumulation point (D_1). The idealized limits of the spin chains are the infinite spin configurations, which countably many site probes tell apart (D_2). A particle that approaches a point from one side only produces one half of a split pair (D_3 , D_4), and a particle that approaches it from both sides at once produces the midpoint (D_7). Light cones that shrink to points, or templates that become infinitely long, give idealized responses supported at a single point (D_5).

The families D_3 and D_4 appear although the hereditarily separable non-metrizable class is empty for observables. There is no contradiction: the closure of the family itself is a split Cantor set, but the idealized responses of the whole observable include the mixtures, and by Theorem 19.1 they form a larger compactum, which is not hereditarily separable.

For the operator theorist. A tree $(\rho_t)_{t \in 2^{<\mathbb{N}}}$ of preparations is a tree of elements of Alice's sort in $\mathcal{M}_C$, and the pointwise closure of its responses consists of the φ -types over $\mathcal{M}_C$ of the elements $[\rho_{t_n}]_{\mathcal{U}}$ of ultrapowers, for sequences of nodes (t_n) and ultrafilters $\mathcal{U}$ on $\mathbb{N}$ (see Variatio 1). Passing to a regular dyadic subtree, by a Ramsey theorem for trees, plays the role that the extraction of indiscernible sequences plays in model theory, and the heptachotomy says that, after this thinning, the configuration of types attached to the tree is one of exactly seven, up to a homeomorphism of the closures that respects the indexing. The class of the

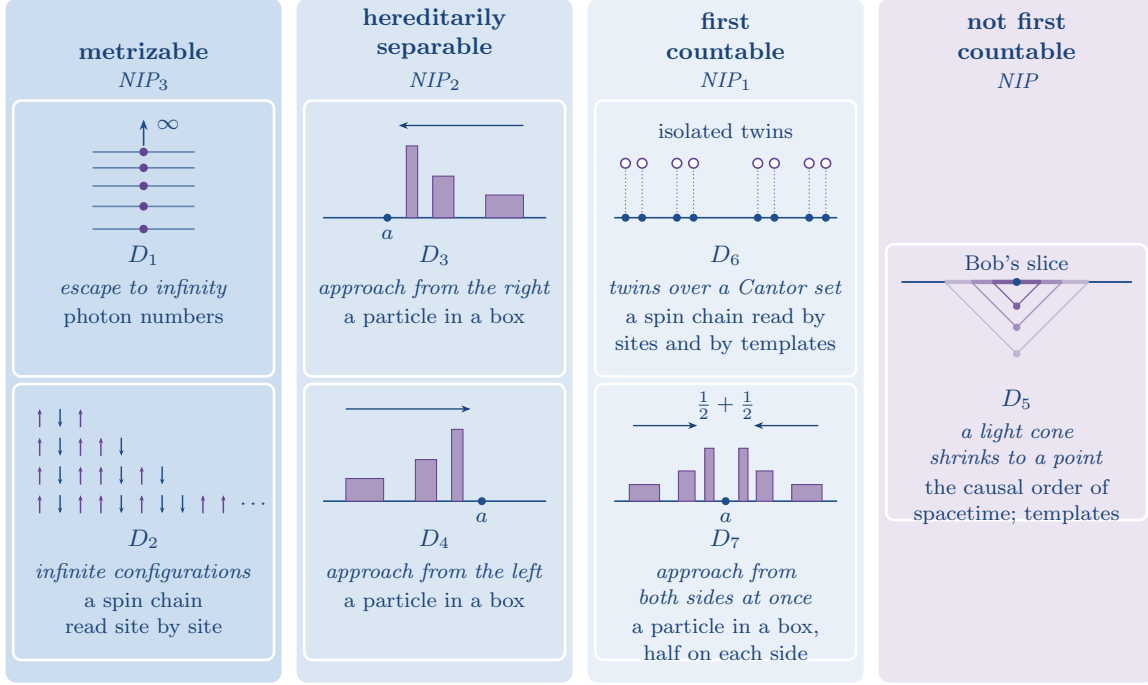


FIGURE 5. The seven minimal families of the heptachotomy, arranged by the class of the Todorčević trichotomy to which their closures belong, each with the physical picture of its realization (see Theorem 15.1): how a tree of preparations approaches its idealized limits.

trichotomy to which the closure belongs, recorded in Table 5, says how many parameters are needed to tell these types apart.

For the logician. In each realization the nodes of the dyadic tree are pure states of Alice's system, and the idealized limits are taken along the branches. For photon numbers the nodes are number states with distinct numbers of photons, so every branch escapes to infinite photon number and all branches share one idealized limit (D_1); for a spin chain read site by site the node t is the chain e_t , the limit along a branch $x \in 2^{\mathbb{N}}$ is the infinite chain x , and its response to the probe at site n is the spin $x(n)$ (D_2). For a particle in the box $[0, 1]$, compared by position with Bob's particle, the nodes are wave functions spread uniformly over short intervals that shrink, along each branch, to a point of a Cantor set, from one side (D_3 , D_4) or half on each side (D_7); the idealized limit answers like a particle at that point, except on probes concentrated at the point itself, where the side of approach determines the answer.

The heptachotomy thus refines the trichotomy into seven ways in which a tree of preparations can fail the double limit while remaining NIP, and physical observables realize all seven. It is not known whether the comparison of positions, which realizes D_7 , also realizes D_6 (see Question 4).

Physical orders in every class

Each of the three classes of the trichotomy that occur for observables contains a physical order, and the interference fringes lie beyond NIP. In this variation we state the result and then treat the examples in turn.

Theorem 16.1 ([DI26b, Theorems 7.8, 7.10, 7.12 and Propositions 7.9, 7.13]).

- (a) *The comparison of two photon numbers is unstable and NIP_3 .*
- (b) *The comparison of two positions on a line, or of two quadratures of light, is NIP_1 but not NIP_2 ; its idealized responses form the Helly space of monotone functions.*
- (c) *The causal order of a $(1+1)$ -dimensional spacetime, “Alice’s event can influence Bob’s detection”, is NIP but not NIP_1 .*
- (d) *The comparison of two phases on a circle, “Bob’s angle is ahead of Alice’s by less than half a turn”, is NIP but not NIP_1 .*
- (e) *The interference fringes of wavenumbers 8^n are not NIP: the closure of their responses contains $\beta\mathbb{N}$.*

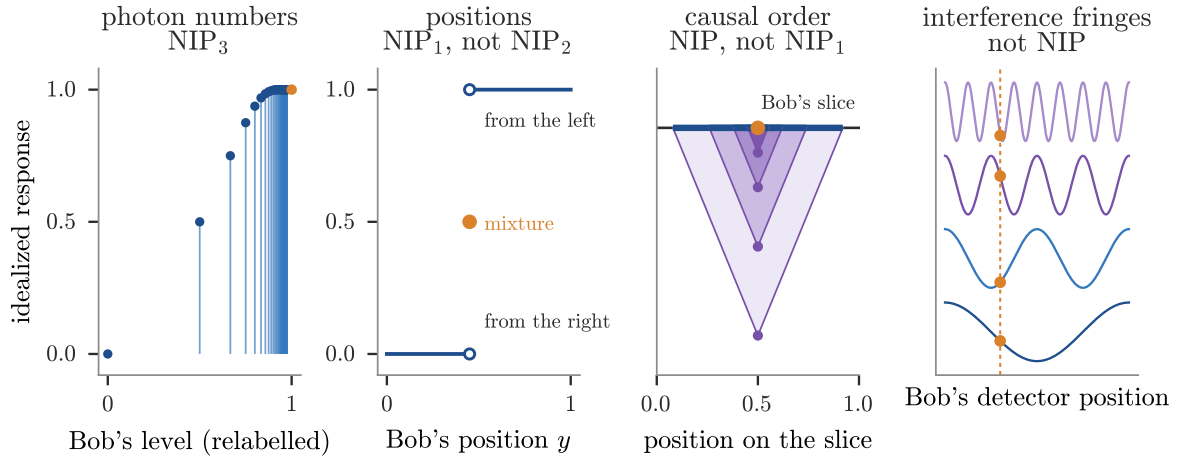


FIGURE 6. Pictures for parts (a), (b), (c) and (e) of Theorem 16.1. Photon numbers: a monotone response on the relabeled spectrum $\{1 - \frac{1}{l+1}\}$, whose points accumulate only at 1. Positions: a particle concentrating at a point from the right or from the left gives two idealized responses that differ only there, and their mixture gives the midpoint. Causal order: events closer and closer to Bob’s time slice, whose light cones meet it in intervals shrinking to a single point. Fringes: at a single detector position, fringes of increasing wavenumber realize a prescribed pattern (shown here 1, 2, 4, 8; the proof uses 8^n).

Photons. The idealized responses of a comparison are monotone, and a monotone function has countably many jumps, each with two sides. For discrete spectra all of them jump only at the countably many points of the spectrum, so countably many probes separate them. Every comparison is moreover uniformly NIP, with bounds that do not depend on the Hilbert spaces: replacing Bob’s distribution by m equally weighted quantile points changes every response

by at most $1/m$, and only polynomially many patterns survive the discretization [DI26b, Proposition 7.3 and Corollary 7.4].

Positions. For continuous spectra every point can be a jump, approached from either side, and the idealized responses form the Helly space.

Light cones. In $1 + 1$ dimensions, an event of Alice influences a detection of Bob exactly when the detection lies between the two points where the light rays through the event cross Bob's time slice. Events whose light cones shrink to a point give an idealized response supported at that point, the picture of the compactum $\hat{A}(2^{\mathbb{N}})$ that witnesses the failure of first countability.

Phases. An antipodal pair of angles on the circle, one approached from behind and the other from ahead, gives an idealized response that equals $\frac{1}{2}$, the response of the uniform state, except at those two points; on a line, a monotone response that is constant away from finitely many points is constant.

Fringes. The response $\frac{1}{2}(1 + \cos 8^n x)$ of a photon prepared with wavenumber 8^n to a detector at position x depends only on $8^n x$ modulo 2π , that is, on the base-8 digits of $x/2\pi$ from the n th on, and these digits can be chosen to realize every finite pattern.

For the operator theorist. Each observable of Theorem 16.1 reads a relation of a classical structure, and the types of that structure display the classes. In the structure of the causal order, with sorts $[0, 1]^2$ and $[0, 1]$ and the relation $a \leq y \leq b$, the element $[(c - 1/n, c + 1/n)]_{\mathcal{U}}$ of the ultrapower, an interval of infinitesimal length around c , has over the real points the φ -type $y \mapsto 1_{\{c\}}(y)$; these types, one for each c , are isolated from one another and accumulate at the zero function, which therefore has no countable base of neighborhoods, the configuration of $\hat{A}(2^{\mathbb{N}})$. For the comparison of positions the corresponding types are the split pairs of Variatio 12, and the idealized responses of the observable, which include those of mixtures, are the monotone functions of the Helly space. In the quantum structure $\mathcal{M}_C$ these types appear as the values of idealized responses at Bob's idealized probes concentrated at points, with Bob idealized first (see Variatio 13).

For the logician. In a $(1 + 1)$ -dimensional spacetime with coordinates (t, ξ) and light speed 1, a signal emitted at an event reaches only the events inside its future light cone, so the event (t, ξ) with $t \leq 0$ can affect a detector at the point y of the slice $t = 0$ exactly when $\xi + t \leq y \leq \xi - t$. The causal-order observable asks this question of an event of Alice and a detector position of Bob, both described by quantum states, and answers yes with the probability that the two positions satisfy the inequality. Events closer and closer to the slice have light cones that meet it in intervals shrinking to a point c , and their idealized limit influences only a detector at exactly c ; since no density operator gives positive probability to a single point, this idealized response is seen only by idealized probes, and continuum many such points cannot be watched by countably many probes. For the phases the corresponding feature is that a circle, unlike a line, carries an idealized response equal to $\frac{1}{2}$ except at two antipodal points.

Physical orders thus occupy the three classes that occur for observables, and the fringes lie beyond NIP. Each of these observables fails the double limit, and its class in the trichotomy measures how far it fails.

Orders of limits in phase transitions

A liquid has a sharp boiling point only in the limit of infinitely many molecules, and a magnet retains its magnetization at zero field only in infinite volume. Phase transitions are failures of the interchange of two limits, as we recall below for the Ising model, and in this variation we show that the classes of the trichotomy detect them.

The Ising model is the basic model of a ferromagnet. On a finite box Λ of $\mathbb{Z}^d$ with N sites, a configuration $s \in \{-1, 1\}^\Lambda$ has energy $E(s) = -J \sum_{x \sim y} s_x s_y - h \sum_x s_x$, where the first sum runs over neighboring sites, $J > 0$ is the coupling and h the applied field; at inverse temperature β the configurations are weighted by $\exp(-\beta \text{energy})$, and the magnetization $m_N(h)$ is the expected mean spin. The weighted configurations form the *Gibbs state*, in which s has probability $Z_N^{-1} e^{-\beta E(s)}$, with normalizing constant $Z_N = \sum_s e^{-\beta E(s)}$, the *partition function*. The *spontaneous magnetization* is $m^* = \lim_{h \downarrow 0} \lim_N m_N(h)$, along boxes exhausting $\mathbb{Z}^d$; for $d \geq 2$ there is a critical β_c with $m^* > 0$ for $\beta > \beta_c$, the *ordered phase*, and $m^* = 0$ for $\beta < \beta_c$. The *pure phases* are the infinite-volume limits of the Gibbs states as $h \downarrow 0$ and as $h \uparrow 0$, with mean spin $\pm m^*$. Spontaneous magnetization is the standard failure of the interchange of limits: the magnetization of a ferromagnet in a field h , taken to infinite volume and then to zero field, is nonzero below the critical temperature, while the other order gives zero. To obtain a bipartite family, let Alice's preparations be finite samples of N spins with a built-in field offset a , and let Bob's probes be applied fields h ; the response is the magnetization $m_N(h - a)$. For the Ising model in dimension $d \geq 2$ the family is stable exactly when the spontaneous magnetization vanishes, and in the ordered phase it is NIP₁ but not NIP₂: the plus and minus phases give two idealized responses that differ at a single point, a *split pair*, and their midpoints form an uncountable discrete set [DI26b, Theorem 9.2]. Figure 7 shows the split pair in the mean-field (Curie–Weiss) model, where the magnetization of N spins can be computed exactly.

For the operator theorist. The magnetizations $m_N(h - a)$ form a formula in the sense of Variatio 1, with Alice's samples, indexed by size and offset, and Bob's fields as the two sorts, and the idealized responses are its φ -types realized in ultrapowers. In the ordered phase, two of them, the plus and the minus phase at offset a , take the values m^* and $-m^*$ at the parameter $h = a$ and agree at every other parameter; they form a split pair, like the types just above and just below a real number in the rational order (see Variatio 12). Each response is monotone in h , so the family is NIP, and the types that take the midpoint value 0 at $h = a$, one for each offset a , form an uncountable discrete set, which places the family in the class $\text{NIP}_1 \setminus \text{NIP}_2$; in the disordered phase the idealized responses are continuous and the family is stable.

For the logician. The Ising model describes a magnet by a spin $s_x = \pm 1$ at each site of a lattice, and thermal equilibrium at inverse temperature β by the Gibbs state; the magnetization $m_N(h)$ is the expected mean spin, an odd nondecreasing smooth function of the applied field h for each finite N . The double limit in question compares $\lim_{h \downarrow 0} \lim_N m_N(h) = m^*$ with $\lim_N \lim_{h \downarrow 0} m_N(h) = 0$: a finite magnet in zero field has no preferred direction, while an infinite magnet in a vanishing positive field keeps a positive magnetization. In the bipartite family, Alice's preparation is a finite sample whose field is shifted by the offset a and Bob's probe is the applied field; an idealized preparation of Alice is an infinite sample, and at the field $h = a$ it responds with m^* or $-m^*$ according to the side from which the field is

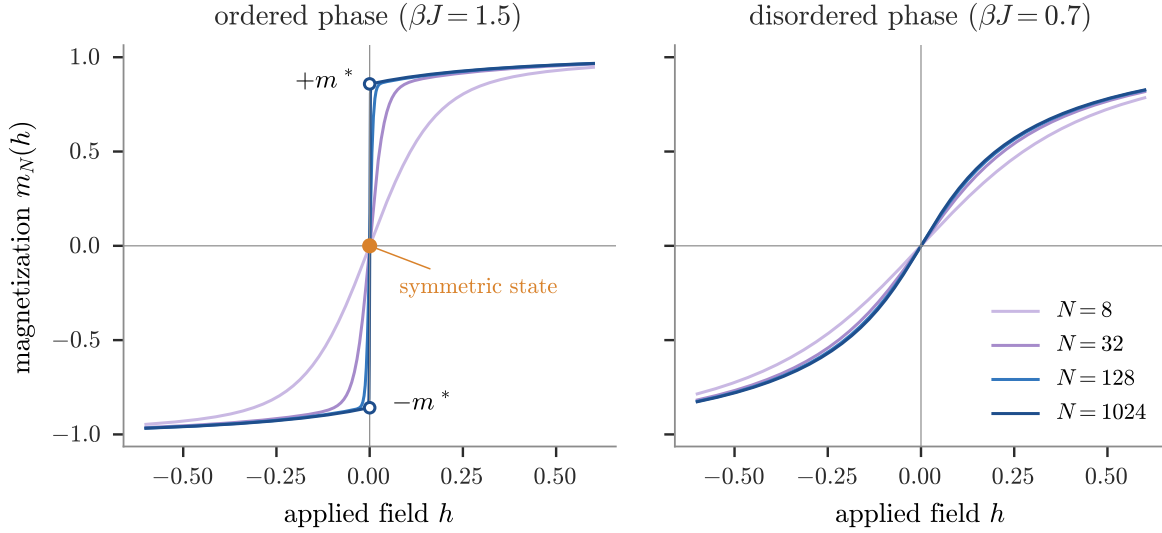


FIGURE 7. Magnetization curves of the Curie-Weiss model for increasing numbers N of spins, with the field h in units of $k_B T$. In the ordered phase the curves steepen into a jump at $h = 0$: approaching zero field from either side gives the two values $\pm m^*$, the split pair, and the symmetric state gives their midpoint 0. In the disordered phase the curves stay smooth.

approached. These two responses are the two pure phases, and their equal mixture, whose mean spin is 0, gives the midpoint.

The trichotomy thus does more than detect a phase transition: the ordered phase places the magnetization in the class $\text{NIP}_1 \setminus \text{NIP}_2$, through the split pair of the two pure phases, and the disordered phase leaves it stable.

Variatio 18. A 1 CLAV. CANONE ALLA SESTA

A quantum ferromagnet in the classification

A magnetic field perpendicular to the axis along which the spins of an Ising magnet align makes the model quantum. In this variation we carry the results of the previous one over to the quantum model.

In a transverse field the Ising Hamiltonian is $H_N = -J \sum \sigma_x^z \sigma_y^z - \Gamma \sum \sigma_x^x$, where each site x carries a qubit, σ_x^z and σ_x^x are the Pauli matrices $\text{diag}(1, -1)$ and $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ acting on the qubit at x (see Variatio 22), the first sum runs over neighboring sites, and $J > 0$ and $\Gamma > 0$. This Hamiltonian no longer commutes with the magnetization. The classification of Variatio 17 uses the differentiability of the pressure away from zero field, which for classical ferromagnets follows from the theorem of Lee and Yang [LY52]. The *pressure* is $p(h) = \lim_N N^{-1} \log Z_N(h)$; it is convex, and where it is differentiable the infinite-volume magnetization is $p'(h)$. The theorem of Lee and Yang states that for a classical Ising ferromagnet on a finite set, the partition function, as a function of a complex field h , vanishes only where $\text{Re } h = 0$. For $\Gamma = 0$ the Hamiltonian is diagonal in the basis of configurations and is the classical Ising energy; for $\Gamma > 0$ it is not, and the Gibbs state is the density operator $e^{-\beta H_N} / \text{Tr } e^{-\beta H_N}$.

Theorem 18.1 ([DI26b, Propositions 9.4, 9.5 and Theorem 9.6]). *For complex h , put $Z_N(h) = \text{Tr } \exp(-\beta H_N + h \sum_x \sigma_x^z)$. Then we have $Z_N(h) \neq 0$ whenever $\text{Re } h \neq 0$, the infinite-volume*

pressure $\lim_N N^{-1} \log Z_N(h)$ is real analytic for real $h \neq 0$, and the magnetization with offsets is stable when the spontaneous magnetization vanishes and NIP_1 but not NIP_2 when it is positive.

The idea is to approximate $\exp(-\beta H_N)$ by the formula of Lie and Trotter, which turns the quantum model into a classical ferromagnet in one more dimension, to apply the classical theorem of Lee and Yang to the approximants, and to pass the zero-free half-planes to the limit by the theorem of Hurwitz. Earlier Lee–Yang theorems for quantum spin systems exclude zeros only in a smaller region (see the historical remarks); we have not found the statement for the whole half-plane in the literature, but its proof combines standard tools and it may be known to experts; the inequality of Griffiths, Hurst and Sherman offers another route to the differentiability.

The formula of Lie and Trotter is $e^{A+B} = \lim_n (e^{A/n} e^{B/n})^n$ for matrices A and B . Applied to $-\beta H_N + h \sum_x \sigma_x^z$, it writes $Z_N(h)$ as a limit of sums over configurations of n stacked copies of the lattice; each factor $e^{\beta \Gamma \sigma_x^x / n}$ couples neighboring copies ferromagnetically, since $\langle s | e^{\varepsilon \sigma^x} | s' \rangle = c e^{K s s'}$ with $K = -\frac{1}{2} \log \tanh \varepsilon > 0$. The result is a classical Ising ferromagnet on the lattice times a cycle of length n , to which the theorem of Lee and Yang applies.

Figure 8 shows the theorem on a small ring of spins.

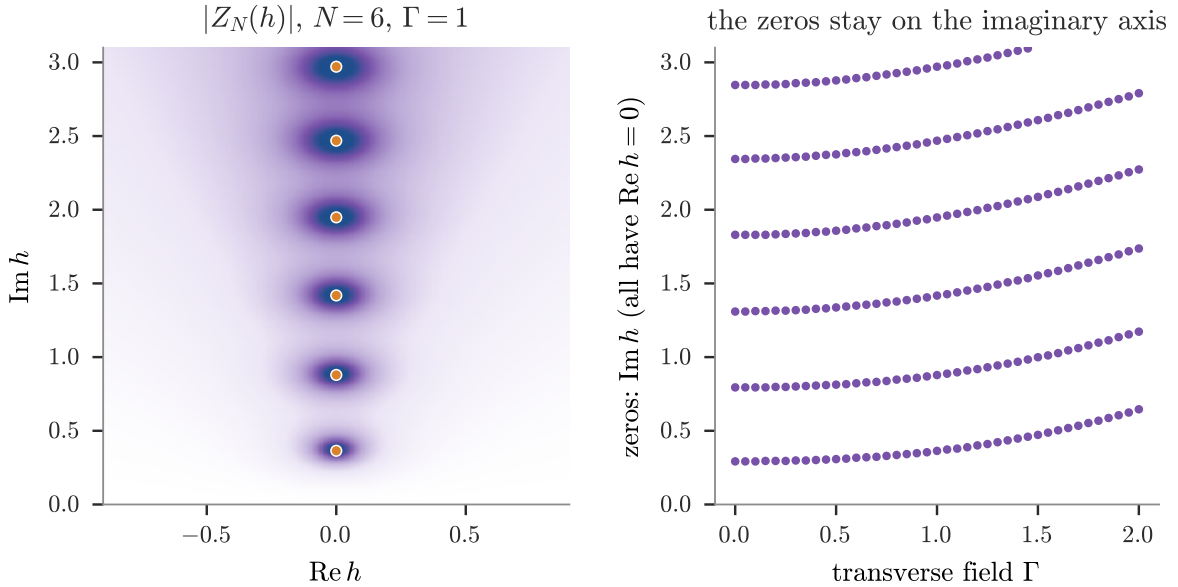


FIGURE 8. Zeros of the partition function $Z_N(h)$ of the Ising model in a transverse field on a ring of $N = 6$ spins, with $\beta = J = 1$, computed by exact diagonalization. Left: the shading is $\log(|Z_N(h)|/Z_N(\text{Re } h))$ for $\Gamma = 1$; the dark spots are the zeros, and all of them lie on the imaginary axis. Right: the zeros as the transverse field grows from the classical model, $\Gamma = 0$; they move, but none leaves the axis.

For the operator theorist. The classification of the family uses two properties of the responses $h \mapsto m_N(h - a)$ and of their limits, and nothing else about the Hamiltonian. First, each response is nondecreasing in h , so no two responses are shattered by the fields: if one

field makes the first response low and the second high, no other field reverses the pattern, and the family has no independence property. Second, each idealized response is continuous except at the single parameter $h = a$, which Theorem 18.1 supplies through the analyticity of the pressure; the split pair, and with it the class $\text{NIP}_1 \setminus \text{NIP}_2$, then follows as in Variatio 17.

For the logician. In the transverse-field model each site carries a qubit, the term $-J\sigma_x^z\sigma_y^z$ favors neighbors aligned along the z axis, and the term $-\Gamma\sigma_x^x$ favors at each site the state $(|0\rangle + |1\rangle)/\sqrt{2}$, which is not an eigenvector of σ_x^z ; since the two terms do not commute, the Gibbs state is a density operator that is not diagonal in the basis of spin configurations. The magnetization is $N^{-1}\partial_h \log Z_N(h)$, the expected value of $N^{-1}\sum_x \sigma_x^z$ in the density operator $\exp(-\beta H_N + h \sum_x \sigma_x^z)/Z_N(h)$, and in infinite volume it is the derivative of the pressure. A phase transition at a field h is a point at which the pressure fails to be analytic; Theorem 18.1 excludes such points away from zero field, so the only jump of the magnetization is the one at zero field, where the two pure phases meet.

When the spontaneous magnetization is positive, the quantum magnet fails the double limit exactly as the classical one does, and for the same reason: a convex pressure, real analytic away from zero field.

Variatio 19. A 1 CLAV.

Convexity forces metrizability

A split pair consists of two idealized responses that differ at a single point, one approached from the left and one from the right. The equal mixture of the two is a third response, their midpoint. In this variation we describe the second of the three effects of mixing announced in the introduction. In Theorem 19.1 below, we show that, for NIP observables, hereditary separability of the idealized responses implies metrizability, and the trichotomy has three classes.

Theorem 19.1 ([DI26b, Theorems 6.6 and 6.9]). *A separable convex Rosenthal compactum that is hereditarily separable is metrizable. Consequently, an NIP observable is NIP_3 , or NIP_1 but not NIP_2 , or not NIP_1 : the trichotomy of Todorćević has three classes for observables, witnessed in the last two cases by D_6 or D_7 , and by D_5 , respectively.*

The idea is the following: a hereditarily separable non-metrizable Rosenthal compactum contains a copy of one of the split Cantor sets $\hat{S}_\pm(2^\mathbb{N})$ of Todorćević, whose points come in split pairs. Following limits of such pairs inside the convex set of idealized responses, the proof in [DI26b] produces their midpoints and shows that the midpoints form an uncountable discrete subspace, which is not separable. For the comparison of positions, the resulting space is the Helly space.

For the operator theorist. In $\mathcal{M}_C$, Alice's sort is convex and φ_C is affine in Alice's variable, so the ultrapower contains, with any elements $[\rho_n]$ and $[\rho'_n]$, the element $[\frac{1}{2}(\rho_n + \rho'_n)]$, whose φ -type is the average of theirs; the idealized responses therefore form a convex compact set of φ -types. For a general NIP formula no such convexity is available, and the formula $x < y$ of the rational order has the split interval as its space of φ -types (see Variatio 12). In Theorem 19.1 we show that convexity excludes this class: the averages of the two types of each split pair are new types, each isolated among them by the parameter at which its pair splits, and they form an uncountable discrete set. In particular, the space of probability

measures on the split interval is not hereditarily separable, contrary to a remark of Godefroy [God80].

For the logician. Every preparation procedure can be randomized: choosing with probability $\frac{1}{2}$ between two procedures and forgetting the choice produces the mixture $\frac{1}{2}(\rho + \rho')$, whose expectations are the averages of those of ρ and ρ' . For the comparison of positions, mixing a particle spread just to the left of a point r with one spread just to the right, and idealizing, gives a preparation that answers $\frac{1}{2}$ on Bob's particle concentrated at r , with Bob idealized first, and agrees with both one-sided limits on probes concentrated away from r ; there is one such preparation for each point of the line. These preparations form an uncountable discrete set of idealized responses, so no quantum observable has the split interval as its space of idealized responses, and the comparison of positions, the closest physical counterpart of the rational order, is NIP_1 but not NIP_2 , with the Helly space as its space of idealized responses (see Theorem 16.1).

In Part I mixtures stood in for idealized preparations. Here they add idealized responses of their own, the midpoints, and the midpoints decide the class.

Part III. Simplicity

Variatio 20. A 2 CLAV.

Simplicity

Each dividing line forbids a combinatorial configuration. Stability forbids the order property and NIP forbids the independence property; simplicity forbids a tree in which every node has infinitely many successors, no k of them consistent together, while every branch is consistent. In this variation we state the theorem that characterizes simplicity. In this part simplicity is first studied in two first-order structures, the random graph and the Pauli group; observables reappear in Variatio 28, where simplicity is studied in Keisler's real-valued logic [Kei23].

A theory is *complete* if it is the theory of a structure. A *monster model* of T is a model $\mathbb{M}$ that is κ -saturated and strongly κ -homogeneous for a very large cardinal κ : every type over a set of fewer than κ elements, a *small* set, is realized in $\mathbb{M}$, and two small tuples with the same type are conjugate by an automorphism of $\mathbb{M}$. The ultrapower of Variatio 1 has the first property for countable sets. We write $\text{tp}(a/B)$ for the type of a over B . Let $\mathbb{M}$ be a monster model of a complete first-order theory T , and let $\varphi(x, y)$ be a formula. A sequence (b_j) is *k-inconsistent* for φ , where $k \geq 2$ is an integer, if no k of the formulas $\varphi(x, b_j)$ have a common solution. The formula has the *tree property* if there are an integer $k \geq 2$ and a tree $(b_s)_{s \in \mathbb{N}^{<\mathbb{N}}}$ whose branches are consistent and in which, at every node, the immediate successors of the node form a k -inconsistent sequence. The formula has the *tree property of the second kind* if there are an integer $k \geq 2$ and an array $(b_{ij})_{i,j \in \mathbb{N}}$ whose rows are all k -inconsistent and in which every choice of one entry per row is consistent. A theory is *simple* if no formula has the tree property [She80]. Stable theories are simple, and a simple theory without the independence property is stable [She90, She80], so the simple unstable theories are exactly the simple theories with the independence property.

Simplicity has a structural characterization, due to Kim and Pillay, which is the usual means of proving that a theory is simple. Kim and Pillay proved that if a ternary relation $A \downarrow_C B$ on small subsets of $\mathbb{M}$ is invariant under automorphisms, symmetric, monotone ($A \downarrow_C BD$ implies $A \downarrow_C B$) and transitive ($A \downarrow_C B$ and $A \downarrow_{CB} D$ imply $A \downarrow_C BD$), has finite character ($A \downarrow_C B$ if and only if $a \downarrow_C b$ for all finite $a \subseteq A$ and $b \subseteq B$)

and local character (for each finite a and each B there is $B_0 \subseteq B$ with $|B_0| \leq |T|$ and $a \downarrow_{B_0} B$), and satisfies extension (for all a and $C \subseteq B$ there is a' with $\text{tp}(a'/C) = \text{tp}(a/C)$ and $a' \downarrow_C B$) and the *independence theorem over models*, then the theory T is simple and the relation is nonforking independence [KP97]. Stable theories have the same kind of characterization, due to Harnik and Harrington [HH84], with *stationarity over models*, the uniqueness of nonforking extensions of a type over a model, in place of the independence theorem; stationarity implies the independence theorem. Uniqueness is thus what separates stability from simplicity. The independence theorem is the substantial axiom: if M is a model, $B \downarrow_M D$, and two realizations a_1 and a_2 of the same type over M satisfy $a_1 \downarrow_M B$ and $a_2 \downarrow_M D$, then some a realizes $\text{tp}(a_1/MB) \cup \text{tp}(a_2/MD)$ with $a \downarrow_M BD$. Over a model, extension is witnessed by the coheirs of Variatio 1; in every theory, the two orders of the iterated ultrapower produce nonforking amalgams, in a stable theory they produce the only one (see Variatio 6), and in a simple theory other nonforking amalgams exist, and the independence theorem governs how they combine.

For the operator theorist. Simplicity is a property of a theory, hence of all elementary extensions of a structure at once, and its configurations are sought in the monster model, an elementary extension far larger than an ultrapower over $\mathbb{N}$ in which every type over a small set is realized. In the language of Variatio 1, a stable theory has over a model exactly one independent amalgam of two types, whereas a simple theory may have many nonforking amalgams, and the theorem of Kim and Pillay replaces uniqueness by the independence theorem: nonforking extensions of one type over a model to two sets independent over the model can be realized together. The tree property is the obstruction: a tree with consistent branches and k -inconsistent siblings yields types that fork over every subset of size at most $|T|$ of their domain, so that nonforking independence loses its local character.

For the logician. The structures of Part III come from the measurement of Pauli strings (see Variatio 22), self-adjoint unitaries on n qubits whose eigenvalues are ± 1 , so that each is a measurement with two outcomes. Two observables can be measured jointly, with a joint probability distribution of outcomes in every state, exactly when they commute, and any two Pauli strings either commute or anticommute; the relation between them that matters for measurement is therefore anticommutation. In quantum error correction an unknown error, itself a Pauli string, is probed by measuring commuting Pauli strings, each outcome recording whether the error anticommutes with the string measured, and the question of this variation, whether partial information about an unknown element can be combined, becomes a question about these outcomes.

Simplicity is not a property of the closure of a family of responses, as stability and NIP are. Its characterization is an amalgamation property, and the double limit reappears only when preparations are mixed.

Variatio 21. CANONE ALLA SETTIMA

The Rado graph

If each pair of points of a countable set is joined by an edge independently with probability $\frac{1}{2}$, the resulting graph is, with probability one, isomorphic to a fixed graph, by a theorem of Erdős and Rényi [ER63]. It is the first of the two basic structures of this part, and the example with which Shelah motivated simple theories. In this variation we show that its theory is simple and unstable, has the independence property, and has forking given by intersection.

The finite graphs form a Fraïssé class, since two graphs that share an induced subgraph can be glued along it without new edges. The Fraïssé limit is the Rado graph (V, E) , the unique countable graph with the *extension property*: for all finite disjoint sets P and Q of vertices, some vertex outside $P \cup Q$ is adjacent to every vertex of P and to no vertex of Q . It is the random graph of the preceding paragraph. Its theory is denoted T_{ind} [She80]. The relation defined by $A \downarrow_C B$ if and only if $A \cap B \subseteq C$ satisfies the axioms of Kim and Pillay [DI26c, Theorem 3.1]. The two substantive axioms can be checked by hand, and we do so, since they show how little a vertex knows.

Forking. Let $C \subseteq B$ be small sets of vertices of a monster model, and let v be a vertex. If $v \notin B$, the type of v over B says only which vertices of B are adjacent to v , and that v differs from each of them; every finite part of it is realized by vertices outside any given finite set, and the type does not fork over C . If $v \in B \setminus C$, the formula $x = v$ divides over C : the type of v over C is not algebraic, so there is a sequence $v = v_0, v_1, v_2, \dots$ of distinct vertices, indiscernible over C , and no vertex equals two of them. Hence the type $\text{tp}(v/B)$ forks over C exactly when $v \in B \setminus C$.

The independence theorem. Let M be a model, let B and D be sets with $B \cap D \subseteq M$, and let a_1 and a_2 realize the same type over M , with $a_1 \downarrow_M B$ and $a_2 \downarrow_M D$. If $a_1 \in M$, then we have $a_2 = a_1$, and the element $a = a_1$ will do. Otherwise $a_1 \notin MB$ and $a_2 \notin MD$, and the two types prescribe, for a new vertex, adjacency to the vertices of MB and of MD . The prescriptions agree on M , because a_1 and a_2 have the same type over M , and they concern disjoint sets outside M , because $B \cap D \subseteq M$. Hence their union is consistent, and saturation, the extension property of the monster model, realizes it by a vertex a outside MBD ; then we have $a \downarrow_M BD$. The edges prescribed by the two types can only meet over M , where they agree.

The extension property also yields the instability of the theory. If the vertices $b_1, a_1, b_2, a_2, \dots$ are chosen in this order by the extension property, one can arrange that a_i is adjacent to b_j exactly when $i < j$: a ladder, so the theory is unstable. Similarly, for distinct vertices $a_1, \dots, a_n$ and every set $S \subseteq \{1, \dots, n\}$, the extension property supplies a vertex adjacent to a_i exactly when $i \in S$: the independence property. Nevertheless, independent information can be combined: a vertex v is independent of B over C exactly when $v \notin B \setminus C$, so the type of a vertex forks only when the vertex itself is named.

For the operator theorist. Over a set B of vertices, the type of a vertex $v \notin B$ is determined by the set of vertices of B adjacent to v ; by the extension property every subset of a finite B occurs in V , and every subset of a countable B occurs in the ultrapower V^u , which is $\aleph_1$ -saturated (see Variatio 1). The independence property is this freedom to prescribe adjacency to a sequence of vertices. Nonforking independence is the relation $A \cap B \subseteq C$, so the type of a vertex forks only when the vertex is named, and the independence theorem amalgamates two types over MB and over MD that agree over a model M because their prescriptions of edges overlap only on M .

For the logician. The Rado graph enters measurement through its diagonal observable $C_E = \sum_{(a,b) \in E} |a\rangle\langle a| \otimes |b\rangle\langle b|$ on $\ell^2(V) \otimes \ell^2(V)$ (see Variatio 14): Alice's register holds a vertex, Bob's register holds a vertex, and the measurement answers yes when the two are adjacent. Since C_E is diagonal, a general preparation ρ of Alice answers every probe as the probability distribution $a \mapsto \langle \rho a, a \rangle$ on vertices does, so superpositions of vertices behave as random vertices. The extension property says that some actual preparation of Bob answers yes to any prescribed finite set of Alice's vertices and no to any other finite set, and in

Variatio 28 we determine what the random vertices of Alice do to the simplicity of this measurement.

Every finite pattern occurs in the random graph, so the double limit fails in it; nevertheless, independent information over a model can be amalgamated.

Variatio 22. A 1 CLAV. ALLA BREVE

The Pauli group

The second basic structure of this part comes from quantum computation. In this variation we show that the Pauli group modulo phases, with its commutation relation, is a Fraïssé limit whose theory is simple and unstable. A *qubit* is the Hilbert space $\mathbb{C}^2$, with the Pauli matrices

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad Y = iXZ.$$

A *Pauli string* on n qubits is a tensor product of n factors, each equal to I , X , Y or Z , such as $XZZXI = X \otimes Z \otimes Z \otimes X \otimes I$. Pauli strings are the basic measurements of quantum information, and they are labeled, up to sign, by vectors of $\mathbb{F}_2^{2n}$: the string with factors $X^{x_k} Z^{z_k}$, made Hermitian by the appropriate power of i , has label (x, z) , and we write P_v for the Hermitian Pauli string with label v . The product of two strings has, up to a phase, the sum of the labels. The crucial fact is that two Pauli strings either commute or anticommute,

$$P_v P_w = (-1)^{\omega(v, w)} P_w P_v, \quad \omega((x, z), (x', z')) = x \cdot z' + z \cdot x',$$

and that ω is a nondegenerate alternating bilinear form on $\mathbb{F}_2^{2n}$. On countably many qubits we let $V = \bigoplus_{k \in \mathbb{N}} \mathbb{F}_2^2$, with the same form; its elements label the Pauli strings that act nontrivially on finitely many qubits.

Figure 9 draws the commutation pattern of the fifteen nontrivial Pauli strings on two qubits. Three pairwise commuting strings whose labels add up to zero form a triangle; there are fifteen of them, and any two commuting strings lie on exactly one. This configuration is the generalized quadrangle of order two, known in quantum information as the *doily*.

Stabilizer codes in brief. We recall what the logician needs. A quantum computer stores information in qubits; in the standard error model, errors are Pauli strings, and a code that corrects a set of Pauli strings corrects every error in their linear span [Got97]. A *stabilizer group* on n qubits is a group $\mathcal{G}$ of pairwise commuting Pauli strings that does not contain $-I$. Its common eigenspace for the eigenvalue 1 is the *code space*; if $\mathcal{G}$ has $n - k$ independent generators, the code space has dimension 2^k and stores k qubits [Got97]. In labels, the group $\mathcal{G}$ corresponds, up to signs, to an isotropic subspace $S \subseteq \mathbb{F}_2^{2n}$, a subspace on which ω vanishes. An error is a Pauli string e ; measuring the generators $g_1, \dots, g_{n-k}$ of $\mathcal{G}$ on the corrupted state reveals the *syndrome* $(\omega(g_1, e), \dots, \omega(g_{n-k}, e))$ without disturbing the stored information: if ψ lies in the code space, then we have $g_i(e\psi) = (-1)^{\omega(g_i, e)} e\psi$, so measuring g_i , that is, the projections $\frac{1}{2}(I \pm g_i)$, gives $(-1)^{\omega(g_i, e)}$ with certainty and leaves $e\psi$ unchanged, and the decoder infers a correction from it. An error with zero syndrome is undetectable. It is harmless if it lies in $\mathcal{G}$, since it then acts trivially on the code space, and it is a *logical error* otherwise. A *logical operator* is a Pauli string that commutes with every element of $\mathcal{G}$ without lying in $\mathcal{G}$ up to phase; it maps the code space to itself and acts on it nontrivially. The *weight* of a Pauli string is the number of its factors different from I , and the *distance* of the code is the least weight of a logical error. A *gate* is a unitary on one or two qubits, tensored with the identity on the others, and a *circuit* is a finite product of gates. A *Clifford*

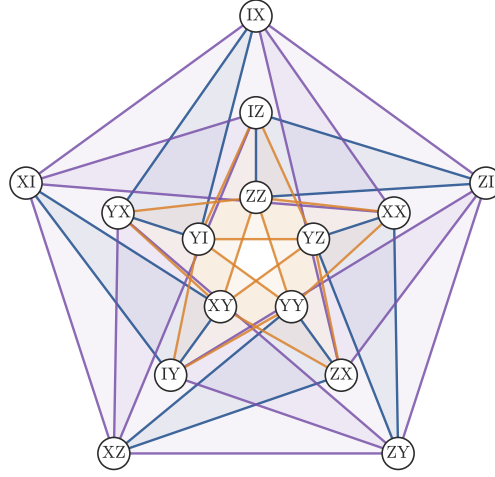


FIGURE 9. The fifteen nontrivial Pauli strings on two qubits. Two strings commute exactly when they lie on a common triangle, and each triangle is a set of three commuting strings whose product is the identity up to a phase. The picture has the five-fold symmetry of a Clifford operation of order five; the colors mark the three orbits of triangles under it.

unitary is a unitary U that maps Pauli strings to Pauli strings under conjugation, up to a phase; on labels, conjugation by U is a linear map that preserves ω .

The structure (V, ω) is the Fraïssé limit of the finite spaces with an alternating form over $\mathbb{F}_2$; the proof, and the model theory of (V, ω) , are in Variatio 24.

Theorem 22.1 ([DI26c, Theorems 3.1 and 4.2]). *The theories of the Rado graph and of (V, ω) are simple and unstable. In (V, ω) , nonforking independence is linear independence over the base: the relation $A \perp_C B$ holds if and only if $\text{span}(AC) \cap \text{span}(BC) = \text{span}(C)$.*

The idea of the proof for (V, ω) is a gluing of linear functionals. A type over a set D records the linear relations of an element over $\text{span}(D)$ and the values of ω against $\text{span}(D)$, which form a linear functional. In the independence theorem, the two functionals prescribed on $\text{span}(MB)$ and on $\text{span}(MD)$ agree on M , and the intersection $\text{span}(MB) \cap \text{span}(MD)$ equals M because $B \perp_M D$; hence they glue to one functional on $\text{span}(MBD)$, which a new element realizes. The independence property comes from nondegeneracy: for linearly independent $a_1, \dots, a_n$, every pattern of values $\omega(a_i, b)$ is realized by some b .

For the operator theorist. The structure (V, ω) is an $\mathbb{F}_2$ -vector space with an alternating form, and its ultrapower V^u is again such a space, an elementary extension in which idealized labels live. The type over a set D of an element $e \notin \text{span}(D)$ is determined by the linear functional $\omega(\cdot, e)$ on $\text{span}(D)$, and every functional occurs, by nondegeneracy when D is finite and by $\aleph_1$ -saturation of the ultrapower when D is countable; this freedom is the independence property. Theorem 22.1 identifies nonforking independence with linear independence over the base, and the independence theorem with the gluing of two linear functionals that agree on a common subspace.

For the logician. On a qubit, X exchanges the basis vectors $|0\rangle$ and $|1\rangle$ (a bit flip), Z multiplies $|1\rangle$ by -1 (a phase flip), and Y does both, up to a phase; a Pauli string applies

such an operation to each of n qubits. Pauli strings serve both as measurements, with outcomes ± 1 , and as errors, the operations that noise applies to a register; every operator on n qubits is a linear combination of them, which is why a code that corrects a set of Pauli errors corrects every error in their span. As an error, acting on density operators by $\rho \mapsto P\rho P^*$, a Pauli string is determined by its label in $\mathbb{F}_2^{2n}$, since phases cancel, and the form ω records which pairs of strings can be measured jointly: those with $\omega(v, w) = 0$.

The Rado graph embeds in (V, ω) (see Variatio 27). Like the Rado graph, (V, ω) is simple and unstable and has the independence property, and its nonforking independence is the gluing of linear functionals over a common base.

Variatio 23. A 2 CLAV.

The dictionary of simplicity

Three notions of model theory (Fraïssé limits, forking and the independence theorem) are at work in quantum error correction under other names: in this variation we set out the third dictionary (see Table 6), whose physical side was introduced in Variatio 22, and present the running example. Each of the next three variations states a theorem, proves it, and reads it twice. A single example runs through all four: the five-qubit code of Laflamme, Miquel, Paz and Zurek [LMPZ96] and of Bennett, DiVincenzo, Smolin and Wootters [BDSW96], the smallest code that corrects an arbitrary error on one qubit.

Model theory	Quantum error correction	
the structure (V, ω)	Pauli strings on countably many qubits, with commutation	Variatio 22
the age of (V, ω) , the finite structures that embed in it	commutation matrices of finite sets of Pauli strings	Proposition 24.1
ultrahomogeneity	Clifford circuits; encoding circuits of codes	Theorem 24.2
algebraic closure	the group generated by a set of Pauli strings	Variatio 22
the type of an error over a set	its syndrome, and whether it lies in the group generated	Theorem 25.1
a forking extension	the error becomes a product of known operators	Theorem 25.1
nonforking with zero syndrome	an undetectable logical error	Corollary 25.2
the independence theorem	syndromes of independent checks can be prescribed independently	Proposition 26.1

TABLE 6. The dictionary of simplicity: the Pauli group as a structure.

Example 23.1. The five-qubit code has the stabilizer generators

$$g_1 = XZZXI, \quad g_2 = IXZZX, \quad g_3 = XIXZZ, \quad g_4 = ZXIXZ,$$

and stores one qubit, with logical operators $\bar{X} = XXXXX$ and $\bar{Z} = ZZZZZ$. Each of the fifteen errors that act as X , Y or Z on a single qubit has its own nonzero syndrome, and

there are exactly fifteen nonzero syndromes of four bits: the code is *perfect*. For instance, the error $XIIII$ has syndrome $(0, 0, 0, 1)$, the error $IIZII$ has syndrome $(0, 0, 1, 0)$, and the error $YIIII$ has syndrome $(1, 0, 1, 1)$.

Variatio 24. A 1 CLAV. CANONE ALL'OTTAVA

Fraïssé limits and Clifford circuits

Let a symmetric matrix of zeros and ones with zeros on the diagonal be given. In Proposition 24.1 and Theorem 24.2 below, we show that some family of Pauli strings anticommutes exactly where the matrix has a one, and that any two linearly independent families with the same pattern are carried to each other by a Clifford circuit. The two readings coincide: for the logician, the Pauli group is a Fraïssé limit; for the physicist, its symmetries are circuits.

In the language with $+$, 0 and a relation for $\omega(x, y) = 1$, the finite spaces with an alternating form amalgamate freely, every finite one embeds in a nondegenerate one, and the theorem of Witt extends isometries between subspaces of a finite nondegenerate space. A back-and-forth argument then shows that (V, ω) is the Fraïssé limit of the finite spaces with an alternating form over $\mathbb{F}_2$. Its theory is ω -categorical and eliminates quantifiers, algebraic closure is linear span, and it is one of the basic smoothly approximable structures of Cherlin and Hrushovski [CH03]. A theory is ω -categorical if all its countable models are isomorphic; by the theorem of Ryll-Nardzewski, this holds exactly when, for each n , the n -tuples of every model, for instance of $(V, \omega)^{\mathcal{U}}$, realize only finitely many types. *Quantifier elimination* says that every formula is equivalent to one without quantifiers; for (V, ω) , two tuples, in V or in $V^{\mathcal{U}}$, have the same type as soon as they satisfy the same linear relations and the same values of ω . An element a is *algebraic* over A if $\text{tp}(a/A)$ has finitely many realizations in the monster model, and the *algebraic closure* of A is the set of such elements. An ω -categorical structure is *smoothly approximable* if it is the union of a chain of finite substructures F_n such that any two tuples of F_n with the same type are conjugate by an automorphism that maps F_n onto itself; for (V, ω) , take for F_n the labels of the first n qubits. The next two results make these facts physical.

Proposition 24.1. *If M is a symmetric $k \times k$ matrix over $\mathbb{F}_2$ with zero diagonal, then there are Pauli strings $P_1, \dots, P_k$ on at most k qubits such that P_i and P_j anticommute exactly when $M_{ij} = 1$.*

Proof. The matrix M defines an alternating form on $\mathbb{F}_2^k$. Let R be its radical and choose a complement W of R ; the form is nondegenerate on W , which therefore has a symplectic basis of r pairs, where $2r = \dim W$. Send the r pairs to the labels of X and Z on the first r qubits, and a basis of R to the labels of X on $\dim R$ further qubits. This linear map preserves the form, and the images of the standard basis vectors of $\mathbb{F}_2^k$ are the labels of the required strings, on $r + \dim R \leq k$ qubits. $\square$

Theorem 24.2. *Let W and W' be subspaces of $\mathbb{F}_2^{2n}$ and let $f: W \rightarrow W'$ be a linear bijection with $\omega(f(v), f(w)) = \omega(v, w)$ for all $v, w \in W$. Then there is a Clifford unitary U on n qubits with $UP_v U^* = \pm P_{f(v)}$ for all $v \in W$. If, moreover, $v_1, \dots, v_m$ is a basis of W , then the unitary U can be chosen with $UP_{v_i} U^* = P_{f(v_i)}$ for all $i \leq m$.*

Proof. By the theorem of Witt [Art57], the map f extends to a linear bijection T of $\mathbb{F}_2^{2n}$ that preserves ω . Every such T is induced by conjugation by a Clifford unitary [CRSS98, DDM03]; let U be one. Since $UP_v U^*$ is a Pauli string up to a phase, Hermitian and with square I , the phase is a sign. For the second statement, let $\varepsilon_i \in \{\pm 1\}$ be the sign with $UP_{v_i} U^* = \varepsilon_i P_{f(v_i)}$. The vectors $f(v_i)$ are linearly independent, so by nondegeneracy there is $b \in \mathbb{F}_2^{2n}$

with $(-1)^{\omega(f(v_i),b)} = \varepsilon_i$ for all i . Conjugation by P_b multiplies $P_{f(v_i)}$ by $(-1)^{\omega(f(v_i),b)}$, so the Clifford unitary $P_b U$ has the required property. $\square$

Corollary 24.3. *Any two stabilizer groups on n qubits with the same number of independent generators are conjugate by a Clifford unitary. In particular, every code that stores k qubits in n has an encoding circuit made of Clifford gates, which maps the trivial code with generators $Z_1, \dots, Z_{n-k}$ onto it.*

Proof. Let $g_1, \dots, g_m$ and $g'_1, \dots, g'_m$ be independent generators, with labels v_i and v'_i . Both families span isotropic subspaces, so the linear map $v_i \mapsto v'_i$ preserves ω , and Theorem 24.2 gives a Clifford unitary with $UP_{v_i}U^* = \pm P_{v'_i}$. The signs of the generators g_i and g'_i are then matched by a Pauli correction, as in the proof of the theorem. $\square$

On countably many qubits, every finite-dimensional subspace of V lies in the labels of the first n qubits for some n . Applying Theorem 24.2 there, and extending by the identity on the remaining qubits, shows that every isometry between finite-dimensional subspaces of (V, ω) extends to an automorphism of (V, ω) implemented by a Clifford circuit on finitely many qubits.

For the operator theorist. Proposition 24.1 identifies the age of (V, ω) , the class of finite structures that embed in it, with the class of finite commutation patterns, and Theorem 24.2 is ultrahomogeneity: every isomorphism between finite substructures extends to an automorphism, here implemented by a Clifford circuit. By quantifier elimination, the type of a tuple of labels, in V or in any ultrapower $V^{\mathcal{U}}$, is its pattern of linear relations and commutations, and for each length there are only finitely many such patterns, which is ω -categoricity. Corollary 24.3 is the case of isotropic subspaces: two families of independent generators of stabilizer groups of the same size have the same type, hence are conjugate by an automorphism, up to signs that a Pauli correction adjusts, and the encoding circuit of a code is that automorphism.

For the logician. The Clifford unitaries are, up to phases, the products of three kinds of gates: the Hadamard gate $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, which exchanges X and Z under conjugation; the phase gate $\text{diag}(1, i)$, which sends X to Y and fixes Z ; and the controlled-NOT gate on two qubits, which flips the second qubit when the first is in $|1\rangle$. Conjugation by a Clifford unitary maps Pauli measurements to Pauli measurements and Pauli errors to Pauli errors, which is why these gates are the basic operations of stabilizer error correction. An *encoding circuit* for a code that stores k qubits in n is a unitary that maps the code space of the trivial code, in which the first $n - k$ qubits are fixed in $|0\rangle$ and the last k carry the information, onto the code space; Corollary 24.3 provides one, built from these gates, for every stabilizer code.

Variatio 25. A 2 CLAV. ADAGIO

Forking and logical errors

Some errors of a stabilizer code commute with every check, and therefore leave every bit of the syndrome at zero: they are undetectable, and the decoder does not correct them. Some undetectable errors are harmless, others destroy the stored qubit, and the syndrome does not distinguish the two kinds.

In this variation we separate them by forking (see Variatio 6). In $V^{\mathcal{U}}$, for instance, if the labels e_n are pairwise distinct, then the class $[e_n]_{\mathcal{U}}$ does not lie in V , and by Theorem 25.1

below its type over any $B \subseteq V$ does not fork over $\emptyset$. We compute forking for the Pauli group from Theorem 22.1, and then read it in a code.

Theorem 25.1. *Let e be an element of a monster model $\mathbb{M}$ of the theory of (V, ω) , and let $D \subseteq B$ be small subsets of $\mathbb{M}$. Then the type $\text{tp}(e/B)$ forks over D if and only if $e \in \text{span}(B) \setminus \text{span}(D)$.*

Proof. By Theorem 22.1, the type $\text{tp}(e/B)$ does not fork over D exactly when $\text{span}(eD) \cap \text{span}(B) = \text{span}(D)$. If $e \in \text{span}(D)$, then we have $\text{span}(eD) = \text{span}(D)$, and the equality holds. If $e \notin \text{span}(D)$, then we have $\text{span}(eD) = \text{span}(D) \oplus \{0, e\}$, and the equality fails exactly when some vector $e + d$ with $d \in \text{span}(D)$ lies in $\text{span}(B)$, that is, exactly when $e \in \text{span}(B)$. $\square$

The theorem says that commutation data never fork: the type of e over B records the values $\omega(b, e)$ for $b \in B$, and these values can be prescribed freely without forking. Forking happens only when e becomes a product of the operators in B . In a stabilizer code this separates the two kinds of undetectable error.

Corollary 25.2. *Let $\mathcal{G}$ be a stabilizer group with set of labels S , and let e be a nonzero label with zero syndrome. Then the type $\text{tp}(e/S)$ forks over $\emptyset$ if and only if $e \in S$, that is, if and only if e is harmless. Consequently, the distance of the code is the least number of qubits acted on by a nonzero error with zero syndrome whose type over S does not fork over $\emptyset$.*

Proof. Apply Theorem 25.1 with $D = \emptyset$ and $B = S$, noting that $\text{span}(S) = S$ and $\text{span}(\emptyset) = \{0\}$. The second statement is the definition of distance, since a logical error is a nonzero label with zero syndrome outside S . $\square$

In the five-qubit code, the errors $g_1 g_2 = XYIYX$ and $\bar{X} = XXXXX$, up to phases, both have zero syndrome. The first lies in the stabilizer group, so its type over S forks, and it is harmless. The second does not, so its type does not fork, and it flips the stored qubit. The least weight of an error of the second kind is 3, as the count below shows, attained for instance by $\bar{X} g_1 = IYYIX$ up to a phase: the code has distance 3.

We now count the undetectable errors of the five-qubit code. On five qubits there are $4^5 = 1024$ Pauli strings up to phase. The undetectable ones, those with zero syndrome, are the labels on which ω vanishes against the four checks; they form a subspace of dimension $10 - 4 = 6$, sixty-four labels in all. Sixteen of them lie in the stabilizer group. One is the identity, which is no error at all, and whose type over S does not fork over $\emptyset$, since it lies in $\text{span}(\emptyset)$. The other fifteen are strings of weight four; their types over S fork over $\emptyset$, and they are harmless. The remaining forty-eight are logical errors, and their types do not fork over $\emptyset$: thirty of weight three, eighteen of weight five, and none of weight four.

weight of the nonidentity undetectable error	3	4	5
harmless: the type over S forks over $\emptyset$	0	15	0
logical: the type over S does not fork over $\emptyset$	30	0	18

Every undetectable error of weight four is harmless. Every undetectable error of weight three is logical. The distance of the code, three, is the weight at which the nonforking errors begin.

Forking also tracks what is learned as operators are added to the checks. Suppose that the logical operator $\bar{Z}$ is added. The type of $\bar{X}$ over $S \cup \{\bar{Z}\}$ records that $\bar{X}$ anticommutes with $\bar{Z}$, so the error is detected. Its type still does not fork over S , because $\bar{X}$ is not a product of

$\bar{Z}$ and the stabilizers. By Theorem 25.1, the type of $\bar{X}$ over a set $B \supseteq S$ forks over S exactly when $\bar{X}$ lies in $\text{span}(B)$.

Forking can be watched directly in an elementary extension. Take $D = \emptyset$, $B = S$, and the harmless error $e = g_1 g_2$, whose label is $g_1 + g_2$. The formula $x = g_1 + g_2$, with parameters g_1, g_2 , divides over $\emptyset$: on fresh qubits of the structure on countably many qubits there are pairs $(g_1^{(j)}, g_2^{(j)})$, $j \in \mathbb{N}$, with the same commutation pattern as (g_1, g_2) and pairwise disjoint supports, hence with pairwise distinct sums; by Ramsey's theorem and compactness they may be taken indiscernible over $\emptyset$, and no label equals two of the sums. Hence the type $\text{tp}(e/S)$ forks over $\emptyset$, as Corollary 25.2 states. For a logical error e , in Theorem 25.1 we show that no formula of $\text{tp}(e/S)$ divides over $\emptyset$: by quantifier elimination, the type says only that $e \notin \text{span}(S)$ and that $\omega(s, e) = 0$ for $s \in S$, and it is realized by labels on qubits disjoint from any given finite set.

Let us notice that the type of a logical error over S is not stationary, and that this is the decoder's situation: all labels outside $\text{span}(S)$ with zero syndrome have the same type over S , since by quantifier elimination that type is determined by the linear functional $\omega(\cdot, e)$ on $\text{span}(S)$, which is 0. Its nonforking extensions to $S \cup \{\bar{Z}\}$ are at least two: the extension realized by $\bar{X}$, which anticommutes with $\bar{Z}$, and the extension realized by a label supported on a sixth qubit, which commutes with $\bar{Z}$; neither label lies in $\text{span}(S \cup \{\bar{Z}\})$, so neither extension forks. A decoder that learns the values of ω on further labels chooses among nonforking extensions, and the error is identified exactly when its type forks, that is, when it becomes a product of the labels that are known.

For the operator theorist. Forking (see Variatio 6) can be seen directly in (V, ω) . If $e = b_1 + \dots + b_m$ with $b_i \in B$ and $e \notin \text{span}(D)$, then the formula $x = b_1 + \dots + b_m$ divides over D : since the sum is not algebraic over D , an elementary extension contains copies of $(b_1, \dots, b_m)$ that form a sequence indiscernible over D with pairwise distinct sums, and no element equals two of them. Conditions on the values $\omega(b, x)$ for $b \in B$, by contrast, never cause forking, by Theorem 25.1. For example, an element $[e_n]_U$ of the ultrapower with pairwise distinct labels e_n lies outside V , so its type over any subset of V does not fork over $\emptyset$, whatever commutation values it carries.

For the logician. An error on a code with stabilizer group $\mathcal{G}$ is undetectable when it commutes with every element of $\mathcal{G}$, and then it maps the code space to itself. If it lies in $\mathcal{G}$ up to phase, it acts on the code space as a scalar and the stored information is intact; otherwise it acts as a nontrivial unitary on the stored qubits, a logical error, and no measurement of elements of $\mathcal{G}$ tells the two cases apart, since both give the trivial syndrome. Corollary 25.2 reads the distinction as forking: the type of the error over the labels of $\mathcal{G}$ forks exactly when the error is harmless. In the five-qubit code every nonidentity error on at most two qubits has a nonzero syndrome, every error on one qubit has its own syndrome and can therefore be corrected (see Example 23.1), and the distance 3 is the weight of the lightest logical error.

Forking thus draws exactly the distinction that the syndrome cannot: an undetectable error is harmless if and only if its type over the stabilizer group forks. The next variation turns to the independence theorem, the other half of the characterization of Kim and Pillay.

The gluing of syndromes

Suppose that two sets of checks on the same qubits are designed independently, in the sense that their spans meet only in the span of their common checks, and that a syndrome is prescribed for each set, the two agreeing on the common checks. The independence theorem says that types that do not fork can be amalgamated; for the Pauli group, in Proposition 26.1 below we show that a single error produces both syndromes.

Proposition 26.1. *Let A_0 , B and D be subsets of $\mathbb{F}_2^{2n}$ with $\text{span}(BA_0) \cap \text{span}(DA_0) = \text{span}(A_0)$, and let f_1 and f_2 be linear functionals on $\text{span}(BA_0)$ and on $\text{span}(DA_0)$ that agree on $\text{span}(A_0)$. Then there is a label $e \in \mathbb{F}_2^{2n}$ with $\omega(s, e) = f_1(s)$ for $s \in \text{span}(BA_0)$ and $\omega(s, e) = f_2(s)$ for $s \in \text{span}(DA_0)$.*

Proof. Since f_1 and f_2 agree on the intersection of their domains, there is a linear functional f on $\text{span}(BA_0) + \text{span}(DA_0)$ extending both. Extend f to $\mathbb{F}_2^{2n}$. Because ω is nondegenerate, every linear functional on $\mathbb{F}_2^{2n}$ has the form $\omega(\cdot, e)$ for a unique e . $\square$

In the five-qubit code, take $A_0 = \emptyset$, $B = \{g_1, g_2\}$ and $D = \{g_3, g_4\}$. The spans of B and D meet only in 0, so any two bits of syndrome on g_1, g_2 and any two on g_3, g_4 are realized together by one error; by Example 23.1, even by an error on at most one qubit, such as $YIII$ for the syndrome $(1, 0 | 1, 1)$. If D is replaced by $\{g_1g_2\}$, the hypothesis fails, and so does the conclusion: the syndrome bit of g_1g_2 is the sum of those of g_1 and g_2 .

For a finite number of qubits, the information carried by a syndrome is measured by counting instead. If $s_1, \dots, s_m$ are linearly independent labels in $\mathbb{F}_2^{2n}$, then for every syndrome $\tau \in \mathbb{F}_2^m$ exactly 2^{2n-m} labels e satisfy $\omega(s_i, e) = \tau_i$ for all i , since the map $e \mapsto (\omega(s_i, e))_{i \leq m}$ is linear and onto. A uniformly random error has a uniformly random syndrome, which reveals exactly m bits. The count measures codimension, which forking does not measure; it reflects the smooth approximability of (V, ω) by finite structures.

For the operator theorist. The independence theorem of Kim and Pillay (see Variatio 20) concerns two nonforking extensions of one type over a model M , to sets B and D independent over M , and asserts that they have a common nonforking realization. In a stable theory this is automatic, since a type over a model has only one nonforking extension to MBD (see Variatio 6); in a simple unstable theory the two extensions are chosen independently, and the theorem says that they never conflict. Proposition 26.1 is the case of (V, ω) in finite dimension, over the closed set $\text{span}(A_0)$ in place of a model: the two types are the prescribed syndromes f_1 and f_2 , the independence of B and D is the hypothesis on the spans, and the common realization is the error e .

For the logician. A syndrome is the list of outcomes, each ± 1 , of measuring the checks of a code on the corrupted state, and it records for each check whether the error anticommutes with it. Proposition 26.1 says that if two families of checks are independent over their common part, in the sense of the hypothesis on spans, then any two lists of outcomes that agree on the common checks are produced by a single Pauli error, so that no outcome of one family excludes an outcome of the other. In the five-qubit code the outcomes of g_1, g_2 and of g_3, g_4 can be prescribed separately, whereas for the checks g_1, g_2 and g_1g_2 the third outcome is the product of the first two. On n qubits, a family of m independent checks reveals exactly m bits about a uniformly random error, each syndrome being produced by the same number 2^{2n-m} of errors.

The independence theorem is thus to simplicity what uniqueness of the independent product (see Theorem 6.2) is to stability: the property of amalgams that characterizes the class.

Variatio 27. A 2 CLAV. CANONE ALLA NONA

The Rado graph inside the Pauli group

Every graph on a set of qubits determines a quantum state, its graph state. In Proposition 27.1 below, we show by means of graph states that every relation between two finite sets is a pattern of commutation between Pauli strings, and in Variatio 28 that, with idealized strings on one side, so is every relation between countable sets, the Rado graph among them. The Pauli group carries in addition the multiplication of labels, and in Variatio 29 we show that this multiplication determines the effect of mixing.

A *stabilizer state* is the one-dimensional code space of a stabilizer group with n independent generators on n qubits. Let G be a graph on a finite set W of qubits. The *graph state* of G is the stabilizer state with the generators

$$K_v = X_v \prod_{w \sim v} Z_w \quad (v \in W),$$

where X_v and Z_w act on the qubits v and w .

Proposition 27.1.

- (a) *For every finite graph G on W and all $u, v \in W$, the strings X_u and K_v anticommute exactly when u and v are adjacent.*
- (b) *For every relation $\varphi \subseteq A \times B$ between finite sets, there are Pauli strings P_a , $a \in A$, and Q_b , $b \in B$, such that P_a and Q_b anticommute exactly when $(a, b) \in \varphi$.*

Proof. (a) The string X_u commutes with every X factor of K_v , and with every Z factor on a qubit other than u ; it anticommutes with Z_u , which is a factor of K_v exactly when $u \sim v$. (b) Place the elements of A on distinct qubits, let $P_a = X_a$, and let Q_b be the product of Z_a over the a with $(a, b) \in \varphi$. $\square$

For a finite graph, Bob's strings in part (a) are actual Pauli strings, the generators of its graph state; the passage to infinite graphs, such as the Rado graph, is made in Variatio 28.

For the operator theorist. In the ultrapower $V^{\mathcal{U}}$ an element $[b_n]$ defines on V the linear functional $v \mapsto \lim_{n, \mathcal{U}} \omega(v, b_n)$, its ω -type over V , and every linear functional on V arises in this way, since each of its restrictions to a finite-dimensional subspace is realized in V by nondegeneracy. Applied to the flips X_a on a countable set of qubits, this shows that every relation between a countable set A and a countable set of idealized labels is a pattern of commutation, which is how the Rado graph lives inside the Pauli group (see Variatio 28). The two theories nevertheless differ: algebraic closure is trivial in the Rado graph and is linear span in (V, ω) , and accordingly a vertex forks only when it is named (see Variatio 21), while a label forks when it becomes a linear combination of given labels (see Theorem 25.1).

For the logician. A graph state is the unique state, up to a phase, left unchanged by every K_v ; it is prepared by placing every qubit in $(|0\rangle + |1\rangle)/\sqrt{2}$ and applying, for each edge $\{v, w\}$, the controlled- Z gate, which multiplies $|1\rangle_v |1\rangle_w$ by -1 and fixes the other three basis vectors. A bit flip X_u applied to a graph state is detected by measuring the stabilizers K_v : by Proposition 27.1(a), the outcome is -1 exactly for the neighbors v of u , so the syndrome of a single flip is a row of the adjacency matrix. For the Rado graph on countably many qubits

each K_v would act on infinitely many qubits, which is why Bob's strings become idealized in Variatio 28, while the flips X_u remain actual.

The Pauli group thus carries, beyond the commutation patterns that it shares with the Rado graph, the linear structure of its labels; in Variatio 29 we show that this structure decides whether the loss of simplicity under mixing is already visible among actual preparations.

Variatio 28. A 2 CLAV.

Mixing destroys simplicity

In Theorem 28.1 below we describe the third effect of mixing announced in the introduction: simplicity is not preserved under the passage from pure preparations to mixtures. The idea is a comparison of two averages: the difference between the responses to the uniform mixtures on two complementary halves of a block of labels measures how a probe weighs the two halves, and no probe gives a large difference for two different halvings of the same block.

Let $\varphi: A \times B \rightarrow \{0, 1\}$ be a relation between countable sets of labels, and consider its diagonal observable (see Variatio 14)

$$C_\varphi = \sum_{a \in A, b \in B} \varphi(a, b) |a\rangle\langle a| \otimes |b\rangle\langle b| \quad \text{on } \ell^2(A) \otimes \ell^2(B) :$$

Alice's register holds a label a , Bob's register holds a label b , and the observable reads $\varphi(a, b)$. A mixture of Alice's preparations is a probability measure μ on A , and her response to a probe ν is the probability $R(\mu, \nu)$ that the observable reads 1. The probes form the space $\mathcal{P}(K_\varphi)$ of Radon probability measures on the closure K_φ of the functions $\varphi(\cdot, b)$ in $\{0, 1\}^A$, the idealized labels of Bob [DI26c, Lemma 6.1]. The observable is thus a two-sorted structure $\mathcal{S}_\varphi$ of real-valued logic, with the mixtures of Alice and the probes of Bob as sorts and the response R as predicate, and we call it *simple* if its theory in Keisler's real-valued logic [Kei23] is simple in the sense of Ben Yaacov [BY03]: no formula $\theta(x, y)$ has the tree property, that is, there are no $\varepsilon > 0$, k and tree of parameters such that the conditions $\theta(x, b) = 0$ along each branch are jointly satisfiable while no k siblings satisfy $\theta(x, b) \leq \varepsilon$ together. A formula is *affine* if it is built with the connectives $u(t_1, \dots, t_n) = c_0 + \sum_i c_i t_i$ and the quantifiers. In an elementary extension, for instance in $\mathcal{S}_\varphi^u$, the new preparations are idealized preparations of Alice.

The commutation observable of the Pauli group is the diagonal observable of ω , and the Rado observable is the diagonal observable of E . Both have the independence property already on pure labels. On countably many qubits, Proposition 27.1(b) extends to countable sets A and B with Bob's strings replaced by idealized ones: the linear functional on V that takes the value 1 at X_a exactly for $(a, b) \in \varphi$ is a point of the compact group V^* of all linear functionals, which is the space of idealized labels of Bob for the commutation observable [DI26c, Lemma 6.3]. In particular, the Rado graph and the random bipartite graph are commutation patterns between single-qubit flips and idealized Pauli strings, while for a finite graph Bob's strings are actual Pauli strings, the generators of its graph state.

Theorem 28.1 ([DI26c, Theorem 7.1]). *Let C_φ be a diagonal observable whose pure responses have the independence property: there are distinct labels $a_0, a_1, \dots$ such that every finite pattern of values $\varphi(a_n, b)$ is realized by some label b . Then the affine formula $R(y_1, x) - R(y_2, x)$ has the tree property of the second kind in the theory of $\mathcal{S}_\varphi$. In particular, a simple diagonal observable has no independence property.*

Proof. Let N be given and choose t with $2^t > N$. Divide $N2^t$ of the labels a_n into N blocks $S_0, \dots, S_{N-1}$, identify each block with $\mathbb{F}_2^t$, choose distinct nonzero $w_0, \dots, w_{N-1} \in \mathbb{F}_2^t$, and let $T_{ij} = \{v \in S_i : w_j \cdot v = 1\}$, a half of the block S_i ; two different halves in a block meet in a quarter of it. Let $\mu_{ij}^\pm$ be the uniform mixtures on T_{ij} and on $S_i \setminus T_{ij}$. For a probe with probabilities q_a of answering 1 at the label a , and with $m = 2^t$, we have

$$R(\mu_{ij}^+, \nu) - R(\mu_{ij}^-, \nu) = \frac{2}{m} \left(\sum_{a \in T_{ij}} q_a - \sum_{a \in S_i \setminus T_{ij}} q_a \right) \leq 1,$$

and for $j \neq j'$ the sum of the two values in row i is $\frac{4}{m} (\sum_{T_{ij} \cap T_{ij'}} q_a - \sum_{S_i \setminus (T_{ij} \cup T_{ij'})} q_a) \leq 1$, so the two values are never both greater than $1/2$. For any choice of one j in each row, the independence property gives a label b that answers 1 exactly on the chosen halves, and at b every chosen value is 1. These finite arrays exist for every N , and compactness yields an infinite one in $\mathcal{S}_\varphi^u$: by the theorem of Loś the classes of the finite arrays form an infinite array whose rows are 2-inconsistent at level $\frac{1}{2}$ and each of whose finite paths is satisfiable, and $\aleph_1$ -saturation realizes every infinite path. $\square$

The theorem is the counterpart, for independent preparations, of the theorem of Ben Yaacov that the randomization of a theory (Keisler's real-valued theory of random elements $\xi: \Omega \rightarrow M$ of a model, on an atomless probability space, with predicates $\mathbb{P}[\theta(\xi_1, \dots, \xi_n)]$) with the independence property has the tree property of the second kind [BY13], and it mirrors Theorem 19.1: in both cases, a class of the classical theory disappears because states can be mixed. The tree of the proof lives in an elementary extension, among idealized preparations. Whether actual preparations already carry it separates the two examples, as we show in the next variation.

For the operator theorist. The structure $\mathcal{S}_\varphi$ has the mixtures of labels as Alice's sort, and in Theorem 28.1 we show that this convex sort carries the tree property of the second kind whenever the labels have the independence property: the parameters of the finite arrays are pairs of actual mixtures, and the theorem of Loś with $\aleph_1$ -saturation turns them into an infinite array in the ultrapower $\mathcal{S}_\varphi^u$. Since a formula with the tree property of the second kind has the tree property, the real-valued theory of $\mathcal{S}_\varphi$ is not simple. The theorem parallels the result of Ben Yaacov on randomizations [BY13]: in both, a structure is replaced by probability measures on it, and the independence property of the original relation produces arrays of averages whose rows are inconsistent.

For the logician. A diagonal observable reads classical data stored in two quantum registers: Alice's register holds a label a as the basis vector $|a\rangle$, Bob's holds b as $|b\rangle$, and the measurement returns $\varphi(a, b)$; on a density operator only the probabilities $\langle \rho a, a \rangle$ of the labels matter, so Alice's preparations are, in effect, random labels. For the commutation observable Alice's label is a Pauli string, a check, and Bob's is an error, and the measurement returns 1 when they anticommute, the syndrome bit of the error at that check. The proof uses only uniform mixtures over halves of blocks of labels, which a laboratory prepares by choosing a label of the half uniformly at random, and the tree property of the second kind is a statement about the probabilities with which such random checks detect a fixed error.

In Part I mixing stood in for idealization; in Part II it supplied the midpoints of split pairs; here it builds a tree.

Variatio 29. A 1 Ô VERO 2 CLAV.

Actual and idealized trees

Theorem 28.1 builds the tree in an elementary extension, among idealized preparations. In this variation we show that for the Pauli group the tree already exists among actual preparations, and that for the Rado graph it does not.

Theorem 29.1 ([DI26c, Theorems 8.1, 9.1 and 9.3]).

- (a) *For the commutation observable, actual preparations witness the tree property of the second kind. Let $(e_{im})_{i,m \in \mathbb{N}}$ be linearly independent labels and let μ_{ij} be the uniform mixture on the coset $e_{ij} + \text{span}(e_{im} : m < j)$. Then we have $R(\mu_{ij}, b) = 1$ exactly when e_{ij} is the first check of block i to anticommute with b ; any choice of one entry in each of finitely many rows is realized by a single label; and for $j \neq j'$ no probe gives both μ_{ij} and $\mu_{ij'}$ a response greater than $2/3$.*
- (b) *For the Rado observable, actual preparations do not witness any tree property: for every continuous connective u and every n , no infinite sequence of instances of the formula*

$$u(R(y_1, x), \dots, R(y_n, x)),$$

with actual mixtures as parameters, is uniformly k -inconsistent, for any k : for some $\varepsilon > 0$, no k instances are simultaneously satisfied up to ε .

The idea of (a) is that a single mixture over a coset tests many checks at once: if the error commutes with the checks before e_{ij} and anticommutes with e_{ij} , every string in the coset anticommutes with the error; if it anticommutes with an earlier check, exactly half of them do. Two indices cannot both describe the first check that anticommutes with the error, and blocks of disjoint checks are independent. The difference between the responses to the uniform mixtures on the coset and on the subgroup is the indicator of a syndrome sector, the projector $|\mathcal{G}|^{-1} \sum_{g \in \mathcal{G}} \chi(g) g$, with χ the character of $\mathcal{G}$ given by the syndrome, computed as an average over a stabilizer group. The idea of (b) is that the responses of the Rado observable depend only on the marginals of a probe, the probabilities q_a , and every vector of marginals is attained. An infinite sequence of conditions on marginals must, after a gliding hump argument, put its weight on disjoint blocks of labels apart from a common part, and one probe built block by block nearly satisfies any k of them.

For the operator theorist. A configuration such as an array witnessing the tree property of the second kind may exist in the structure $\mathcal{S}_\varphi$ itself or only in an elementary extension, such as the ultrapower $\mathcal{S}_\varphi^u$, whose new elements are idealized preparations; in Theorem 29.1 we show that both cases occur, as the ladders of Part I separate stability from uniform stability. For the commutation observable, Bob's idealized labels form the compact group V^* of linear functionals on V (see Variatio 28), Alice's labels are its characters, and averages over cosets recover, by Fourier inversion on V^* , indicators of cylinder sets, which build the array from actual mixtures. For the Rado observable the space of Bob's idealized labels is the whole Cantor cube $\{0, 1\}^V$, responses depend only on the marginals of a probe, and no infinite sequence of instances with actual parameters is uniformly k -inconsistent.

For the logician. Alice's preparation μ_{ij} is the random check obtained by multiplying e_{ij} by a uniformly random product of the earlier checks e_{im} , $m < j$, of block i ; against a fixed error it answers 1 with certainty if e_{ij} is the first check of block i that anticommutes with the error, with probability $\frac{1}{2}$ if an earlier check anticommutes with it, and with probability

0 otherwise. Consequently, the difference between the responses to the uniform mixtures on the coset and on the subgroup is the indicator of a syndrome sector, the set of errors with prescribed outcomes on the checks e_{im} , $m \leq j$. For the Rado observable no mixture of vertices has this effect, and the same tree requires averages over infinitely large blocks of vertices, uniform distributions that exist only as finitely additive limits, that is, as idealized preparations.

The distinction between actual and idealized configurations, which Part I drew for ladders, thus reappears for trees.

Variatio 30. A 1 CLAV. QUODLIBET

The commutation relation

A single relation of quantum mechanics runs through the three parts: a momentum kick and a translation commute up to a phase. Read with bounded kicks, with unbounded kicks and with binary kicks, it gives an example for each part, as we show in Proposition 30.1 below.

For a particle on a line, let T_x translate a wave function by x , so that $(T_x\psi)(q) = \psi(q - x)$, and let K_k give it a momentum kick k , so that $(K_k\psi)(q) = e^{ikq}\psi(q)$. With position Q , multiplication by q , and momentum $P = -i d/dq$, in units with $\hbar = 1$, one has $T_x = e^{-ixP}$ and $K_k = e^{ikQ}$, and $K_k^*PK_k = P + k$: the kick shifts momentum by k . Then we have

$$K_kT_x = e^{ikx}T_xK_k,$$

the commutation relation of Weyl, the bounded form of the canonical commutation relation $QP - PQ = iI$; by the theorem of Stone and von Neumann, it determines (Q, P) up to unitary equivalence for irreducible families. On a qubit, the bit flip X is a translation and Z , multiplication by $(-1)^q$ for $q \in \{0, 1\}$, is a kick, and $ZX = -XZ$: the Pauli group is the group of Weyl over the two-element field, with phase $(-1)^{\omega(v,w)}$ in place of e^{ikx} . The phase e^{ikx} also determines the intensity that Bob observes when Alice's photon, prepared with wavenumber k , reaches his detector at position x : the interference fringe of Variatio 16 is $\frac{1}{2}(1 + \operatorname{Re} e^{ikx})$.

Proposition 30.1. *The commutation phase, read as a bipartite observable in which Alice chooses a kick k and Bob a translation x , the position of his detector, yields three observables according to the range of its parameters, and they lie in three different classes.*

- (a) Bounded parameters: stable. *If Alice's wavenumbers k_n lie in a bounded set, then the fringe observable $C = \sum_n |e_n\rangle\langle e_n| \otimes M_{f_n}$ on $\ell^2(\mathbb{N}) \otimes L^2[0, 2\pi]$, with $f_n(x) = \frac{1}{2}(1 + \cos k_n x)$, is stable.*
- (b) Lacunary parameters: not NIP. *If $k_n = 8^n$, the same observable is not NIP, and its idealized responses contain a copy of $\beta\mathbb{N}$.*
- (c) Binary parameters: simple, and not simple once mixed. *Over the two-element field, on countably many qubits, the commutation relation is the structure (V, ω) , which is simple and unstable; the observable that reads it has the tree property of the second kind, witnessed by actual mixtures of Alice's preparations.*

Proof. (a) The effective observable of a mixture $\sum_n t_n |e_n\rangle\langle e_n|$ is multiplication by $\sum_n t_n f_n$. The functions f_n are uniformly bounded and, since $|f'_n(x)| \leq \frac{1}{2}|k_n|$, equicontinuous; by the theorem of Arzelà and Ascoli they form a relatively compact set in $C[0, 2\pi]$, and so does its closed convex hull. A norm-compact set of effective observables is weakly compact, and the observable is stable by Theorem 5.1(A2). The off-diagonal parts of Alice's states do not

contribute, because C is diagonal in Alice's factor. (b) This is [DI26b, Proposition 7.13]. (c) This is Theorem 22.1 together with Theorem 29.1(a). $\square$

The proposition concerns three restrictions of one relation, not one observable in three classes: the bounded and the unbounded fringes are two observables, and the qubit version is a third. With bounded parameters it satisfies the double limit criterion of Grothendieck; with the wavenumbers 8^n it is not NIP. Over the two-element field it is the arithmetic of error correction, simple in itself and not simple once preparations are mixed.

For the operator theorist. The three regimes are three formulas. With bounded wavenumbers the responses f_n form a norm-compact set of continuous functions, so every idealized response is a continuous function of the detector position and every φ -type realized in the ultrapower is definable, as in Theorem 3.1. With the wavenumbers 8^n the formula has the independence property, since the responses at a single position realize every finite pattern, and the φ -types realized in ultrapowers contain a copy of $\beta\mathbb{N}$ (see Theorem 11.1). Over $\mathbb{F}_2$ the relation is the alternating form of (V, ω) , a standard example of a simple unstable theory, and the mixtures of Alice's preparations carry the tree property of the second kind already in the structure itself (see Theorem 29.1).

For the logician. A translation T_x moves a particle by x and a kick K_k adds k to its momentum; applied in the two orders, they produce the same state up to the factor e^{ikx} , which is invisible for a single state but observable in interference, where two amplitudes are added. In the fringe observable this phase is measured directly: a photon of wavenumber k is detected at the position x with relative intensity $\frac{1}{2}(1 + \cos kx)$, and Bob's choice of x is the choice of a translation. On a qubit the bit flip X is the translation of $\{0, 1\}$ by 1 and the phase flip Z is a kick, so the relation $ZX = -XZ$ is the same relation with phases ± 1 ; on n qubits it becomes the commutation relation ω of Pauli strings, which a stabilizer measurement reads. The proposition follows one physical relation through the three dividing lines, according to whether the parameters range over a bounded set of reals, over the lacunary set $\{8^n\}$, or over the two-element field.

ARIA DA CAPO E FINE

Physics lives on idealized limits. A physicist speaks freely of systems that no laboratory prepares, and the variations took the physicist at that word: the idealized preparations of Alice and the idealized probes of Bob were elements of ultrapowers, and their responses were types. For a single system, idealization remained harmless throughout.

The difficulty began with two systems, and it stayed with them to the end. The staircase failed the double limit from the first page, and so is unstable; the variations placed it further. The comparison of the two energies is NIP, and its idealized responses lie in the metrizable class of the trichotomy.

The agreement of the two orders determined, as the Aria said it would, the operational meaning of an idealized system. Where the orders agreed, idealized preparations were simulated by finite mixtures of actual ones, two idealized systems had a unique independent product, and, in the uniform version, an unknown preparation could be learned (see Theorems 5.1 and 7.1). Where they disagreed, all of this failed.

The theme of this paper was the theorem of Grothendieck, and it was the bass of all thirty variations. In Part I it was heard plainly. In Part II it was broken, and what remained of it

was sequential: every idealized response was still the limit of a sequence of actual ones. In Part III it fell silent, and the amalgamation of independent information took its place.

A logician recognized the staircase as the order property, and the three lines that Shelah drew became three distinctions among measurements. Mixing, the operation by which a quantum observable differs from an arbitrary formula, altered each of them: it realized idealized preparations, it made hereditary separability imply metrizability, and it destroyed simplicity.

Part I studied the observables that satisfy the criterion of Grothendieck, Part II the ways in which it can fail while idealized responses remain sequential limits of actual ones, and Part III the amalgamation of independent information that takes its place in simple theories.

QUESTIONS

Question 1. In classical model theory, a simple theory without the independence property is stable. Is a simple observable stable? By Theorem 28.1, a simple diagonal observable has no independence property, so the question is whether the unstable NIP observables of Part II, such as the energy comparison of the introduction, acquire a tree property once preparations are mixed.

Question 2. Which diagonal observables with the independence property have the tree property of the second kind witnessed by actual preparations? Theorem 29.1 suggests a criterion: whether the closed linear span of the responses contains the indicators of an infinite independent sequence of nested cylinders, with bounded norm.

Question 3. The causal order of a $(1+1)$ -dimensional spacetime is NIP but not NIP_1 . In which class of the trichotomy of Todorćević does the causal order of a spacetime of higher dimension lie?

Question 4. The minimal family D_6 is realized in Part II by a spin chain read both site by site and against templates, and D_7 by the comparison of positions of a particle in a box. Does the comparison of positions also realize D_6 ?

Question 5. For qudits of prime dimension p , the labels form a symplectic space over the field with p elements, and the dictionary of Variatio 23 carries over. For continuous variables, the commutation phase of a translation by x and a momentum kick by k is e^{ikx} , whose real part gives the interference fringe of Part II; for unbounded wavenumbers the fringe is not NIP. Where does the unbounded fringe lie with respect to simplicity?

HISTORICAL REMARKS

Shelah introduced stable theories in 1969 [She69], and the order property and the independence property of formulas in 1971 [She71]; simple theories followed in 1980 [She80], where he named as his motivation the theory T_{ind} of the random graph. The double limit condition also has an analytic history: it is the criterion of Grothendieck for weak compactness of 1952 [Gro52], it is closely related to the regularity of bilinear forms studied by Arens in 1951 [Are51], and it defines the stable Banach spaces of Krivine and Maurey of 1981 [KM81]; Iovino connected it with Shelah's stability and developed the real-valued theory, including the stability spectrum for real-valued structures [Iov99a, Iov99b]. Rosenthal characterized the Banach spaces containing ℓ^1 in 1974 [Ros74], Bourgain, Fremlin and Talagrand studied pointwise compact sets of Baire class one in 1978 [BFT78], and Todorćević proved his trichotomy in 1999 [Tod99]; Argyros, Dodos and Kanellopoulos extracted the heptachotomy from the methods of his paper in 2008 [ADK08]. Theorem 19.1 is a special case of a known result: as Todorćević pointed out to us, hereditarily separable locally connected Rosenthal

compacta are metrizable, by Theorems 5 and 7 of [Tod99] and Lemma 2.4 of Gruenhage [Gru88]. Kim proved the symmetry of forking in simple theories in 1998 [Kim98], and Kim and Pillay the independence theorem and the characterization of simplicity in 1997 [KP97]; Ben Yaacov extended simplicity to continuous logic [BY03]. The random graph was described by Ackermann in 1937 [Ack37] and by Rado in 1964 [Rad64], and Erdős and Rényi proved in 1963 that it is the almost sure countable random graph [ER63]. Fraïssé introduced his limits in 1954 [Fra54], and Ben Yaacov extended them to metric structures [BY15]. Smoothly approximable structures are due to Cherlin and Hrushovski [CH03]. The symplectic description of the Pauli group and of stabilizer codes is due to Gottesman [Got97] and to Calderbank, Rains, Shor and Sloane [CRSS98], and the five-qubit code to Laflamme, Miquel, Paz and Zurek [LMPZ96] and to Bennett, DiVincenzo, Smolin and Wootters [BDSW96]. Keisler introduced randomizations in 1999 [Kei99], and Ben Yaacov proved in 2013 that randomization destroys simplicity in the presence of the independence property [BY13]. Lee and Yang proved their theorem for the Ising ferromagnet in 1952 [LY52]. Asano [Asa70] and Suzuki and Fisher [SF71] extended it to quantum spin systems, the anisotropic Heisenberg ferromagnets whose coupling along the field dominates the transverse couplings; with a transverse field Γ , these results, in the form given by Fröhlich and Rodríguez [FR12, Theorem 5], exclude zeros only for $\operatorname{Re} h > \beta\Gamma$. The approximation of a quantum spin system by a classical Ising system in one more dimension, used in Variatio 18 to reach the whole half-plane, is due to Suzuki [Suz76]. Hong, Ou and Mandel observed their two-photon interference in 1987 [HOM87]. Model theory has met operator algebras and quantum theory along other routes: Farah, Hart and Sherman studied stability in the model theory of operator algebras [FHS13], where it is a property of the theory of an algebra rather than of a single observable; Zilber gave a model-theoretic semantics of the canonical commutation relation [Zil16]; Ben Yaacov showed that randomization preserves NIP [BY09], of which the uniform NIP of comparisons is an elementary quantitative instance; and Hrushovski and Pillay studied invariant measures in NIP theories [HP11].

ACKNOWLEDGMENTS

We thank Stevo Todorčević for pointing out to us that Theorem 19.1 is a special case of a known result, and for the references that establish it.

The form of this paper is that of the Goldberg Variations of Johann Sebastian Bach, published in 1741: an aria, thirty variations in which every third variation up to the twenty-seventh is a canon, at intervals rising from the unison to the ninth, a quodlibet as the thirtieth, and the aria again. The section titles follow Bach's markings. Figure 10 measures this survey against the music: the share of each variation in six performances of the Goldberg Variations, four on the piano and two on the harpsichord, and its share of the words here.

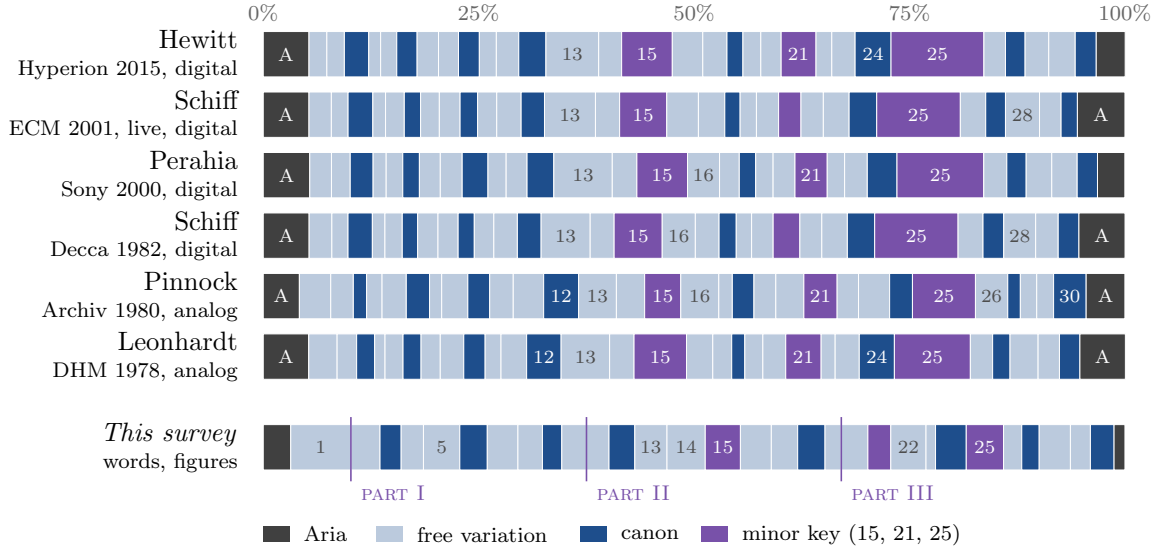


FIGURE 10. The proportions of the Goldberg Variations in performances by Angela Hewitt (piano; Hyperion, 2015), András Schiff (piano; ECM, 2001, live), Murray Perahia (piano; Sony, 2000), András Schiff (piano; Decca, 1982), Trevor Pinnock (harpsichord; Archiv, 1980) and Gustav Leonhardt (harpsichord; Deutsche Harmonia Mundi, 1978), most recent first; the harpsichord recordings are analog and the piano recordings digital, Schiff's of 1982 among the early digital ones. The last bar gives the proportions of this survey by word count, each figure counting as a fixed number of words. Each bar has the same width; the pieces are the Aria (A), Variations 1–30 and the Aria da capo.

REFERENCES

- [Ack37] W. Ackermann, *Die Widerspruchsfreiheit der allgemeinen Mengenlehre*, Math. Ann. **114** (1937), 305–315.
- [ADK08] S. A. Argyros, P. Dodos, and V. Kanellopoulos, *A classification of separable Rosenthal compacta and its applications*, Dissertationes Math. **449** (2008), 1–55.
- [AH14] Hiroshi Ando and Uffe Haagerup, *Ultraproducts of von Neumann algebras*, J. Funct. Anal. **266** (2014), no. 12, 6842–6913.
- [Are51] R. Arens, *The adjoint of a bilinear operation*, Proc. Amer. Math. Soc. **2** (1951), 839–848.
- [Art57] E. Artin, *Geometric Algebra*, Interscience, New York, 1957.
- [Asa70] T. Asano, *Lee–Yang theorem and the Griffiths inequality for the anisotropic Heisenberg ferromagnet*, Phys. Rev. Lett. **24** (1970), 1409–1411.
- [BCWdW01] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, *Quantum fingerprinting*, Phys. Rev. Lett. **87** (2001), 167902.
- [BDSW96] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, *Mixed-state entanglement and quantum error correction*, Phys. Rev. A **54** (1996), 3824–3851.
- [BFT78] J. Bourgain, D. H. Fremlin, and M. Talagrand, *Pointwise compact sets of Baire-measurable functions*, Amer. J. Math. **100** (1978), no. 4, 845–886. MR 509077
- [BY03] I. Ben Yaacov, *Simplicity in compact abstract theories*, J. Math. Log. **3** (2003), 163–191.
- [BY09] ———, *Continuous and random Vapnik–Chervonenkis classes*, Israel J. Math. **173** (2009), 309–333.
- [BY13] Itai Ben Yaacov, *On theories of random variables*, Isr. J. Math. **194** (2013), 957–1012.
- [BY14] ———, *Model theoretic stability and definability of types, after A. Grothendieck*, Bull. Symb. Log. **20** (2014), no. 4, 491–496. MR 3294276

- [BY15] I. Ben Yaacov, *Fraïssé limits of metric structures*, J. Symbolic Logic **80** (2015), 100–115.
- [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.
- [CH03] G. Cherlin and E. Hrushovski, *Finite Structures with Few Types*, Annals of Mathematics Studies, vol. 152, Princeton University Press, 2003.
- [CRSS98] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, *Quantum error correction via codes over $GF(4)$* , IEEE Trans. Inform. Theory **44** (1998), 1369–1387.
- [CS16] A. Chernikov and S. Shelah, *On the number of Dedekind cuts and two-cardinal models of dependent theories*, J. Inst. Math. Jussieu **15** (2016), 771–784.
- [DDM03] J. Dehaene and B. De Moor, *Clifford group, stabilizer states, and linear and quadratic operations over $GF(2)$* , Phys. Rev. A **68** (2003), 042318.
- [DFJP74] W. J. Davis, T. Figiel, W. B. Johnson, and A. Pelcynski, *Factoring weakly compact operators*, J. Functional Analysis **17** (1974), 311–327.
- [DI26a] E. Dueñez and J. Iovino, *Classification of bipartite observables I: stable and uniformly stable observables*, preprint; arXiv:2610.11002, 2026.
- [DI26b] ———, *Classification of bipartite observables II: NIP observables and the Todorčević trichotomy*, preprint; arXiv:2610.12908, 2026.
- [DI26c] ———, *Classification of bipartite observables III: simplicity, mixtures and quantum error correction*, preprint, 2026.
- [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.
- [ER63] P. Erdős and A. Rényi, *Asymmetric graphs*, Acta Math. Acad. Sci. Hungar. **14** (1963), 295–315.
- [FHS13] Ilijas Farah, Bradd Hart, and David Sherman, *Model theory of operator algebras I: Stability*, Bull. Lond. Math. Soc. **45** (2013), no. 4, 825–838.
- [FR12] J. Fröhlich and P.-F. Rodríguez, *Some applications of the Lee–Yang theorem*, J. Math. Phys. **53** (2012), 095218.
- [Fra54] R. Fraïssé, *Sur l’extension aux relations de quelques propriétés des ordres*, Ann. Sci. École Norm. Sup. **71** (1954), 363–388.
- [God80] Gilles Godefroy, *Compacts de Rosenthal*, Pacific J. Math. **91** (1980), no. 2, 293–306. MR 615679
- [Got97] D. Gottesman, *Stabilizer codes and quantum error correction*, Ph.D. thesis, California Institute of Technology, 1997, arXiv:quant-ph/9705052.
- [Gro52] A. Grothendieck, *Critères de compacité dans les espaces fonctionnels généraux*, Amer. J. Math. **74** (1952), 168–186. MR 47313
- [Gro84] Ulrich Groh, *Uniform ergodic theorems for identity preserving Schwarz maps on W^* -algebras*, J. Operator Theory **11** (1984), no. 2, 395–404. MR 86a:46079
- [Gru88] G. Gruenhage, *[title to be confirmed]*, General Topology and its Relations to Modern Analysis and Algebra VI (Proc. Sixth Prague Topological Symposium, 1986) (Z. Frolík, ed.), Res. Exp. Math., vol. 16, Heldermann, Berlin, 1988.
- [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.
- [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.
- [Iov99a] José N. Iovino, *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.
- [Kei99] H. Jerome Keisler, *Randomizing a model*, Adv. Math. **143** (1999), no. 1, 124–158. MR 1680650
- [Kei23] H. Jerome Keisler, *Model theory for real-valued structures*, Beyond First-Order Model Theory, Vol. 2 (José Iovino, ed.), CRC Press, 2023, pp. 43–101.
- [Kim98] B. Kim, *Forking in simple unstable theories*, J. London Math. Soc. (2) **57** (1998), 257–267.
- [KM81] Jean-Louis Krivine and Bernard Maurey, *Espaces de Banach stables*, Israel J. Math. **39** (1981), no. 4, 273–295. MR 83a:46030
- [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 99b:03049
- [Lit88] N. Littlestone, *Learning quickly when irrelevant attributes abound: a new linear-threshold algorithm*, Machine Learning **2** (1988), 285–318.
- [LMPZ96] R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek, *Perfect quantum error correcting code*, Phys. Rev. Lett. **77** (1996), 198–201.
- [LY52] T. D. Lee and C. N. Yang, *Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model*, Phys. Rev. **87** (1952), 410–419.
- [Maz33] S. Mazur, *Über konvexe Mengen in linearen normierten Räumen*, Studia Math. **4** (1933), 70–84.
- [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.
- [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.
- [Rad64] R. Rado, *Universal graphs and universal functions*, Acta Arith. **9** (1964), 331–340.
- [Ray02] Y. Raynaud, *On ultrapowers of non commutative L_p spaces*, To appear, 2002.
- [Ros74] H. P. Rosenthal, *A characterization of Banach spaces containing l^1* , Proc. Nat. Acad. Sci. U.S.A. **71** (1974), 2411–2413. MR 358307
- [Sau72] N. Sauer, *On the density of families of sets*, J. Combinatorial Theory Ser. A **13** (1972), 145–147. MR 307902
- [SF71] M. Suzuki and M. E. Fisher, *Zeros of the partition function for the Heisenberg, ferroelectric, and general Ising models*, J. Math. Phys. **12** (1971), 235–246.
- [She69] S. Shelah, *Stable theories*, Israel J. Math. **7** (1969), 187–202.
- [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
- [She80] S. Shelah, *Simple unstable theories*, Annals of Math. Logic **19** (1980), 177–203.
- [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 91k:03085
- [Sim68] S. Simons, *The iterated limit condition, a Fubini theorem and weak compactness*, Math. Ann. **176** (1968), 87–95.
- [Sim15] Pierre Simon, *Rosenthal compacta and NIP formulas*, Fund. Math. **231** (2015), no. 1, 81–92. MR 3361236
- [Suz76] M. Suzuki, *Relationship between d -dimensional quantal spin systems and $(d + 1)$ -dimensional Ising systems*, Progr. Theoret. Phys. **56** (1976), 1454–1469.
- [Tod99] Stevo Todorčević, *Compact subsets of the first Baire class*, J. Amer. Math. Soc. **12** (1999), no. 4, 1179–1212. MR 1685782
- [Zil16] B. Zilber, *The semantics of the canonical commutation relation*, arXiv:1604.07745, 2016.