GAMES IN THE MATRIX

ILIJAS FARAH

1. Games

Let's start with the simplest interesting mathematical structures. A *simple graph* is a set of vertices some of which are connected by an 'edge' (the technical term is 'adjacent'). What makes a graph simple is that only distinct nodes can be adjacent and that we do not allow multiple edges between any two vertices. We write $G = (V, E)$, where V is the set of vertices and E is the set of connections, called edges. If V has m elements, then there are exactly $\binom{m}{2}$ ways to choose the edges, because for every pair of distinct vertices we have two choices, whether to connect them with an edge or not. To compare two graphs, we will play a game. Given graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ and $n \geq 1$, the game $\Gamma(G_1, G_2, n)$ between two players, Challenger and Duplicator is played in n innings and has the following rules. In each inning Challenger chooses a vertex in either one of the graphs. Duplicator responds by choosing a vertex the other graph. Thus in the j-th inning a vertex a_j in V_1 and a vertex b_j in V_2 are chosen. We add the (unnecessary) requirement that vertices cannot be repeated; the first player who repeats a vertex loses the game. After n innings, Duplicator wins if all i and j satisfy $a_i \mathrel{\mathrm{E}_1} a_j$ if and only if $b_i \mathrel{\mathrm{E}_2} b_j$ (another way to say this is that the map $a_i \mapsto b_i$, for $i \leq n$, is an *isomorphism between the induced subgraphs*); otherwise Challenger wins.

Let's get some easy observations out of the way. If there is an isomorphism Φ between the graphs, then Duplicator can win the game by applying Φ or its inverse to the vertices provided by the Challenger. If both graphs have n vertices, then a game in which Duplicator wins produces an isomorphism between the two graphs.

Each game considered in this note is a finite game of perfect information, meaning that each player has complete knowledge of the history of the game up to their current turn. For example, both chess and go are finite games of perfect information, but in each of these games draw is a possible outcome. Since draw is not a possible outcome in any of the games considered in this note, being a finite game of perfect information implies that one of the players has a winning strategy.[1] A *winning strategy* in $\Gamma(G_1, G_2, n)$ for Duplicator is a function from positions in the game in which it is Duplicator's turn whose values are valid moves in the game, and such that by following this strategy Duplicator wins the game regardless of how Challenger plays. A winning strategy for Challenger is defined analogously. If each of the players had a winning strategy, then pitting these strategies against each other would result in a game in which both players win; contradiction. An example

Date: November 12, 2025.

Partially supported by NSERC.

[1]The question of the existence of a strategy for one of the players (a property called determinacy) for infinite games is the first in a long list of fascinating topics related to our subject that we regrettably do not have time to discuss.

of a winning strategy is provided in the proof of Theorem 1.1 below. In chess and go there are three possible outcomes: (i–ii) One of the players has a winning strategy, or (iii) each of the players has a strategy that assures that they do not lose the game.

We move on to a (somewhat) nontrivial example of a winning strategy for Duplicator. For $m \geq 3$ let C_m be the *cycle graph*, whose vertices are $\{1, \dots, m\}$ and two vertices i, j are adjacent if and only if $|i - j| = 1$ or $\{i, j\} = \{1, m\}$.

Theorem 1.1. *If* $\min(m, k) > 2^{n+1}$*, then Duplicator wins* $\Gamma(C_m, C_k, n)$.

Proof. The *distance*, $d(x, y)$ between two vertices x, y in a graph is the smallest l such that there is a path from x to y of length l in which each step consists of moving to an adjacent vertex.[2] For $j \geq 1$ we say that two vertices x, y in C_m are j-far if $d(x, y) > 2^j$, and that a path is j-long if its length is $> 2^j$. Note that vertices are 0-near if and only if they are adjacent. We can now describe the strategy for the Duplicator. The positions in this game are so similar that there is no harm in pretending that the Challenger always chooses vertices from the first graph. To increase the stakes, we switch to the first-person narrative. In the first inning, we can play any vertex b_1. In the second move, we have vertices a_1 and a_2 in one of the graphs and b_1 in the second. If one of the paths is n-short (meaning, not n-long), then choose b_2 so that $d(b_1, b_2) = d(a_1, a_2)$. The salient point is that, because $\min(m, k) > 2^{n+1}$, the longer path between a_1 and a_2 is n-long, and so is

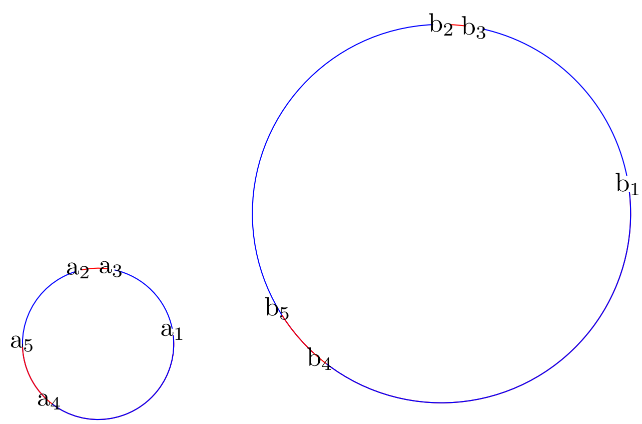

FIGURE 1. The position after five innings. Since m and l are very large, we zoomed out so much that the vertices blended into two continuous circles. The corresponding short (red) arcs are of the same length, while the long (blue) arcs are simply sufficiently long. No matter what vertex the Challenger plays, we can respond by choosing a vertex in the corresponding arc that is either at the right distance or sufficiently far from the corresponding ends.

the longer path between b_1 and b_2. If a_1 and a_2 are n-far, then take b_2 to be any vertex n-far from b_1; this is possible because $\min(m, k) > 2^n$. At this stage we have vertices a_1 and a_2 'splitting' one cycle into two arcs and b_1 and b_2 'splitting' the

[2] If there is no such l then we set $d(x, y) = \infty$.

other cycle into two arcs. Either all of the arcs are sufficiently long, or the short arcs are of the same length.

Our strategy in the $k+1$-st inning for $k \geq 2$ hinges on the following inductive hypothesis holding after the k-th inning.

(IH$_k$) The vertices $a_1, \dots, a_k$ and $b_1, \dots, b_k$ are arranged on their cycles so that for all i and j either $d(a_i, a_j) = d(b_i, b_j)$ or a_i and a_j are $k+1$-far and b_i and b_j are $k+1$-far.

After the k-th inning for $k \geq 2$, Challenger adds a vertex a_{k+1} to one of the arcs, between vertices a_i and a_j for some i, j. If it is added to a short arc, then we can choose b_3 so that sides of the 'triangle'[3] b_i, b_j, b_k are equal to the corresponding sides of the 'triangle' with the vertices a_i, a_j, a_k. If a_k is added to a long arc, but it is $n-1$-near a_i, then we play b_k on the arc b_i, b_j so that $d(b_k, b_i) = d(a_k, a_i)$. Otherwise, play b_k that is k-far from both b_i and b_j. After this move, the inductive hypothesis is satisfied. This shows that we can hold on for at least n moves. When the game is over, b_i and b_j are 0-near if and only if a_i and a_j are 0-near. Since '0-near' is the same as 'adjacent', we win. □

This tells us that sufficiently long cycles are similar. What else is new, you may wonder? What if we choose graphs at random, by fixing the vertex set and then for each pair of vertices flipping a coin to decide whether they are adjacent? If for all j, G_j is a randomly chosen graph with j vertices, then for every n there is $f(n)$ such that if $\min(m, k) \geq f(n)$ then Duplicator wins $\Gamma(G_m, G_k, n)$. This follows from the so-called 0-1 law for random graphs ([8]) and an observation that requires a syntactical definition.[4] .

A first-order statement about graphs is expressed by quantification over the vertices, and it involves the adjacency and equality relations. We will consider only statements that have no free variables; they are called 'sentences' (as opposed to 'formulas').[5] For example (writing $x \mathrel{\mathrm{E}} y$ for 'x and y are adjacent)

$$\psi : (\exists x)(\forall y) y \neq x \to y \mathrel{\mathrm{E}} x$$

asserts that some vertex is adjacent to all other vertices, while

$$\varphi : (\forall x)(\forall y) x \neq y \to (\exists z) z \mathrel{\mathrm{E}} x \wedge \neg(z \mathrel{\mathrm{E}} y)$$

asserts that for any two distinct vertices x, y some vertex is adjacent to x but not to y. A fun little calculation shows that as the number of vertices in a sufficiently random graph converges to ∞, the probability that ψ holds in it converges to 0 and the probability that φ holds in it converges to 1. The 0-1 law asserts that every first-order statement behaves in one of these two ways: in a sufficiently large and sufficiently random graph, it is either almost certainly true or almost certainly false.

But what does this have to do with games? Well, everything. If G_1 satisfies φ and G_2 does not, then the Challenger wins $\Gamma(G_1, G_2, 3)$, as follows. We switch perspective and identify with the winning side. In the first two innings, we choose

---

[3]These triangles are rather flat, but no worries.

[4]A mathematical logician is a mathematician who takes the syntax seriously.

[5]A second-order statement is one in which quantification over arbitrary subsets of the structure is allowed. It is considerably more expressive than the first-order logic. The latter lies in the 'Goldilocks zone' in terms of expressiveness, and in the final section we will see some of the evidence for this.

a_1 and a_2 in G_2 so that no z satisfies $z \mathrel{\mathrm{E}} a_1$ and $\neg(z \mathrel{\mathrm{E}} a_2)$. No matter what the choice of b_1 and b_2 is, in the third inning we can choose b_3 so that $b_3 \mathrel{\mathrm{E}} b_1$ and $\neg(b_3 \mathrel{\mathrm{E}} b_2)$, and the Duplicator is doomed.

For every first-order statement ψ there is n large enough so that the Duplicator loses $\Gamma(G_1, G_2, n)$ if G_1 and G_2 'disagree' on the truth of ψ. The least such n is the *quantifier depth* of ψ; let's just say that the name is well-chosen instead of giving a proper definition. Conversely, if G_1 and G_2 agree on the truth of all first-order statements of quantifier depth n, then Duplicator wins $\Gamma(G_1, G_2, n)$.

These are the Ehrenfeucht–Fraïsse games, affectionately known as the EF-games (see [6, §3.3]). Their variations can be used to compare mathematical structures of all sorts: linear orderings, groups, rings,... In each case, the EF-games provide efficient means for verifying whether two given structures satisfy the same first-order statements. Before moving on, we state another fun fact. It applies to arbitrary structures but for simplicity we stick to graphs.

Theorem 1.2. *If G_m, for $m \in \mathbb{N}$, is any sequence of graphs, then there is a subsequence $G_{m(i)}$, for $i \in \mathbb{N}$, such that for every n there is $k(n)$ such that if $\min(m(i), m(j)) > k(n)$ then Duplicator wins $\Gamma(G_{m(i)}, G_{m(j)}, n)$.*

Thus every sequence of graphs has a subsequence such that any two graphs in it that are sufficiently 'further down the road' cannot be distinguished by a short EF-game. To prove Theorem 1.2, one observes that the set of first-order sentences in the language of graph theory can be enumerated as φ_j, for $j \in \mathbb{N}$. It then suffices to assure that our sequence satisfies the '0-1 law', that all but finitely many $G_{m(i)}$ 'agree' on the truth of every φ_j. This is achieved by a diagonalisation argument.

2. Matrices

During a legendary 1925 vacation in Heligoland, Werner Heisenberg noticed that because of the noncommutativity of matrix multiplication, matrix algebras provide natural setting for quantum mechanics. A few years later, John von Neumann pointed out that the infinite-dimensional analog of $\mathbb{M}_m$ is more suitable for this purpose, and suggested algebras of operators known as II_1 factors as a model. We will see a II_1 factor (or perhaps many II_1 factors—that is the question that we are aiming at) in the final section. Let's first take a closer look at what the matrix algebras are all about.

Consider the rotation ρ_α of the plane by an angle α around the origin, or the (orthogonal) projection p to the line $x = y$ (Fig. (2)). These transformations are implemented by matrices, via

$$p\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) = \begin{pmatrix} \frac{a}{2} + \frac{b}{2} \\ \frac{a}{2} + \frac{b}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix},$$

$$\rho_\alpha\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) = \begin{pmatrix} \cos(\alpha)a - \sin(\alpha)b \\ \sin(\alpha)a + \cos(\alpha)b \end{pmatrix} = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}.$$

In case you haven't seen it yet, the product of a 2×2 matrix and a 2-vector is defined as follows: the first entry of the image of a vector $\begin{pmatrix} a \\ b \end{pmatrix}$ is the dot product of the first row of the matrix with the vector, and its second entry is the dot product of the first row of the matrix with the vector. This works in the higher-dimensional

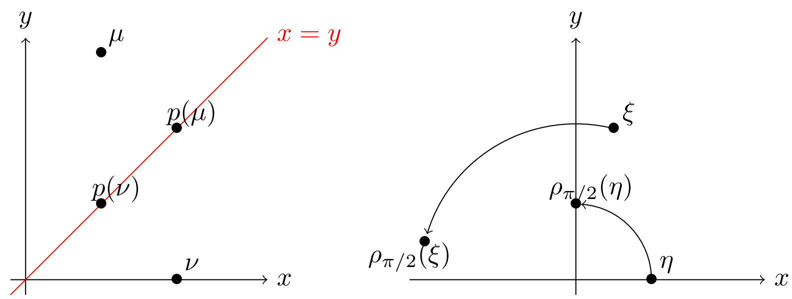

FIGURE 2. The projection to the line $x = y$ and the rotation by $\pi/2$ around the origin as applied to some points randomly chosen in the plane.

context, for example when $n = 3$:

$$\begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^3 c_{1i} a_i \\ \sum_{i=1}^3 c_{2i} a_i \\ \sum_{i=1}^3 c_{3i} a_i. \end{pmatrix}.$$

To multiply two $m \times m$ matrices, a and b, one applies this formula to each column of b and produces the corresponding column of the result, ab.

Fix $m \geq 2$ and let $\mathbb{C}^m$ be the space of m-tuples (called m-vectors) of complex numbers.[6] This is a *vector space* over $\mathbb{C}$; what this means is that we can add vectors pointwise and multiply them by a complex number (called scalar). Some $T\colon \mathbb{C}^m \to \mathbb{C}^m$ is a *linear transformation* if for all vectors v, w and all scalars x, y it satisfies

$$T(xv + yw) = xT(v) + yT(w).$$

Lemma 2.1. *Every linear transformation of $\mathbb{C}^m$ is implemented via multiplication by an $m \times m$ matrix, and multiplication by a fixed $m \times m$ matrix is a linear transformation of $\mathbb{C}^m$. Multiplication of matrices corresponds to composition of linear transformations.*

Thus studying linear transformations of the m-dimensional space is the same as studying the algebra of $m \times m$ matrices; we opt for the latter. We will also need a third operation, the adjoint a^*: to compute it, transpose the matrix and take complex conjugates of its entries. Thus $((a_{ij})_{i,j \leq m})^* = (\overline{a}_{j,i})_{i,j \leq m}$. The adjoint of the matrix that corresponds to a rotation ρ_α is the matrix that implements the rotation $\rho_{-\alpha}$, i.e., its multiplicative inverse. The matrices whose adjoint is their multiplicative inverse are called *unitary*. A matrix that corresponds to an (orthogonal) projection is its own adjoint; such matrices are called *self-adjoint* or *Hermitian*.

[6]Why not the real numbers? It is a long story, but ultimately because we like every polynomial to have a root and we prefer an isometry such as $\rho_{\pi/2}$ to be diagonalisable. However, all that I am about to say makes perfect sense if $\mathbb{C}$ is replaced by $\mathbb{R}$, and a reader who feels more comfortable believing that $x^2 \geq 0$ for all x is encouraged to do so, ignore the adjoint operation defined below, and skip the final section.

By $\mathbb{M}_m$ we denote the algebra of all complex $m \times m$ matrices $a = (a_{ij})_{i,j \leq m}$, equipped with multiplication, addition, adjoint, and multiplication by scalars. (Multiplying a matrix by a scalar appropriately stretches—or shrinks–a given transformation.) By 1_m we denote the *unit matrix* that has 1's down the main diagonal and 0's elsewhere; in other words, $a_{ij} = 1$ if $i = j$ and $a_{ij} = 0$ otherwise. It has the property that $1_m a = a 1_m = a$ for all a in $\mathbb{M}_m$.

2.1. **Let the games begin.** For l, m, and n one could consider the following game, $\Gamma_{\mathrm{ts}}(\mathbb{M}_l, \mathbb{M}_m, n)$. Our two players choose elements from $\mathbb{M}_l$ and $\mathbb{M}_m$ as they did for graphs, and Duplicator wins if the chosen matrices $a_1, \dots, a_n$ and $b_1, \dots, b_n$ 'look the same'. This means that for all i, j, k and all complex numbers y, z we have $a_i a_j = a_k$ if and only if $b_i b_j = b_k$, $y a_i + z a_j = a_k$ if and only if $y b_i + z b_j = b_k$, and $a_i^* = a_j$ if and only if $b_i^* = b_j$.

One could indeed consider this game, but we will not (the subscript 'ts' stands for 'too strict'). Strict equality is fine for graphs and other discrete structures, but when in the matrix there is no need for exactness.[7] Let's make this precise.

The length of an m-vector ξ with coordinates $x_1, \dots, x_m$ is defined as

$$\|\xi\| = \sqrt{\textstyle\sum_{i=1}^m |x_i|^2}.$$

This is justified by the Pythagorean theorem if $m = 2$, supplemented with an induction argument for larger values of m. There are at least two common ways to define the length (called 'norm') of a matrix. First, every $a = (a_{ij})_{i,j \leq m}$ in $\mathbb{M}_m$ is an m^2-vector, and its *normalized Hilbert–Schmidt* norm is

$$\|a\|_{\mathrm{HS}} = \sqrt{\tfrac{1}{m} \textstyle\sum_{i,j=1}^m |a_{ij}|^2}.$$

It is 'normalized' because of the factor $\frac{1}{m}$, assuring that $\|1_m\|_{\mathrm{HS}} = 1$.

The *operator norm* may be slightly less intuitive and it is notoriously difficult to compute for an arbitrary matrix, but it is in some sense even more natural than $\| \cdot \|_{\mathrm{HS}}$. It is defined as

$$\|a\|_{\mathrm{OP}} = \sup\{\|a\xi\| : \|\xi\| = 1\}.$$

This is the 'maximal stretching coefficient': $\|a\|_{\mathrm{OP}} = r$ means that $\|a\xi\| = r\|\xi\|$ for some[8] nonzero ξ and $\|a\xi\| \leq r\|\xi\|$ for all ξ. One has $\|1_m\|_{\mathrm{OP}} = \|1_m\|_{\mathrm{HS}} = 1$ and $\| \cdot \|_{\mathrm{HS}} \leq \| \cdot \|_{\mathrm{OP}}$. On the other hand, any $m \times m$ matrix a that has m in one entry and 0 in all other entries satisfies $\|a\|_{\mathrm{HS}} = 1$ and $\|a\|_{\mathrm{OP}} = m$.

2.2. **Let the correct games begin.** For l, m, n, and $\varepsilon > 0$, we will consider the game $\Gamma_{\mathrm{HS}}(\mathbb{M}_l, \mathbb{M}_m, n, \varepsilon)$ which is played like $\Gamma_{\mathrm{ts}}(\mathbb{M}_l, \mathbb{M}_m, n)$, with two differences:

(P1) Players can choose only matrices x that satisfy $\|x\|_{\mathrm{OP}} \leq 1$.[9]

(P2) Duplicator wins if and only if for all i, j, k and all scalars y, z such that $\max(|y|, |z|) \leq 1$, each of $|\|a_i\|_{\mathrm{HS}} - \|b_i\|_{\mathrm{HS}}|$, $|\|a_i a_j - a_k\|_{\mathrm{HS}} - \|b_i b_j - b_k\|_{\mathrm{HS}}|$, $|\|y a_i + z a_j - a_k\|_{\mathrm{HS}} - \|y b_i + z b_j - b_k\|_{\mathrm{HS}}|$, and $|\|a_i^* - a_j\|_{\mathrm{HS}} - \|b_i^* - b_j\|_{\mathrm{HS}}|$ is no greater than ε.

---

[7]Please don't read much into this sentence.

[8]Do you see why is the sup attained?

[9]The subscript OP is not a typo. The correct ambient for our purposes is the operator unit ball of $\mathbb{M}_m$, regardless of which one of the norms we consider..

The game $\Gamma_{\mathrm{OP}}(\mathbb{M}_l, \mathbb{M}_m, n, \varepsilon)$ is defined analogously, with the payoff in (P2) computed using $\|\cdot\|_{\mathrm{OP}}$ instead of $\|\cdot\|_{\mathrm{HS}}$.

For fixed m and l, as $n \to \infty$ and/or $\varepsilon \to 0$, the game becomes more difficult for the Duplicator. As in the case of graphs, the Duplicator wins $\Gamma_*(\mathbb{M}_l, \mathbb{M}_m, n, \varepsilon)$ (for $* \in \{\mathrm{HS}, \mathrm{OP}\}$) if $\mathbb{M}_l$ and $\mathbb{M}_n$ agree (up to ε) on the values of continuous first-order statements (appropriately defined using one of the two norms, see (1) below) of quantifier depth n.

The time has come to talk of many kinds of matrices. A matrix a is a *projection* if it is self-adjoint ($a = a^*$) and idempotent ($a^2 = a$). The *rank* of a projection is the dimension of its range space $\{a\xi : \xi\}$. A matrix v is a *partial isometry* if v^*v is a projection. This implies that vv^* is also a projection, and ranks of the projections v^*v and vv^* are equal because v is an isometry between their ranges. For example, $v = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ is a partial isometry in $\mathbb{M}_2$ with $vv^* = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $v^*v = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ projections of rank 1. Note that $v^*v + vv^* = 1_2$. If $v \in \mathbb{M}_m$ is a partial isometry such that $v^*v + vv^* = 1_2$, then m is exactly twice the rank of v^*v, hence even. Conversely, if m is even then such partial isometry exists.

Theorem 2.2. *For all sufficiently small $\varepsilon > 0$ and all even m and odd l, Challenger wins $\Gamma_{OP}(\mathbb{M}_l, \mathbb{M}_m, 6, \varepsilon)$.*

Proof. Since m is even, some partial isometry $a_1 \in \mathbb{M}_m$ satisfies $a_1^*a_1 + a_1a_1^* = 1_m$. In the first six innings Challenger plays $a_1, a_2 = a_1^*$, $a_3 = a_1^*a_1$, $a_4 = a_1a_1^*$, $a_5 = a_1^*a_1 + a_1a_1^*$, and $a_6 = 1$[10] Then $a_5 = 1_m$ and since l is odd, Duplicator cannot choose $b_1, \dots b_5$ to satisfy the equalities in (P2) exactly. The trick is that they cannot be satisfied in $\mathbb{M}_l$ even approximately.

If some $b \in \mathbb{M}_m$ satisfies $\|b^*b - (b^*b)^2\|_{\mathrm{OP}} < 1/4$ and $\|bb^* + b^*b - 1_m\|_{\mathrm{OP}} < 1/4$, then some $v \in \mathbb{M}_m$ satisfies $v^*v = (v^*v)^2$, and $v^*v + vv^* = 1_m$. If you know how to diagonalise Hermitian matrices v^*v and vv^*, then this is an exercise. Therefore winning the approximate game is as difficult for the Duplicator as winning the 'on the nose' game, and they lose. □

Just like the discrete case, the conclusion of this theorem can be reformulated as asserting that some first-order statement is evaluated differently at $\mathbb{M}_m$ and $\mathbb{M}_l$. The first-order statement associated with the proof of Theorem 2.2 is the following[11]

$$\psi : \inf_{\|x\|_{\mathrm{OP}} \leq 1} \max(\|x^*x - (x^*x)^2\|_{\mathrm{OP}}, \|x^*x + xx^* - 1\|_{\mathrm{OP}}). \tag{1}$$

The value of ψ in $\mathbb{M}_n$ is 0 if n is even and $> 10^{-10^{10}}$ if n is odd.[12] A variant of the proof of Theorem 2.2 relying on a modification of ψ shows that for every $k \geq 2$, there are n and $\varepsilon > 0$ such that if k divides m but it does not divide l, then Challenger wins $\Gamma_{\mathrm{OP}}(\mathbb{M}_l, \mathbb{M}_m, n, \varepsilon)$.

In other words, the first-order theory of $\mathbb{M}_m$ 'knows' what are the divisors of m. Actually, the first-order theory of $\mathbb{M}_m$ even 'knows' the value of m, but the point

[10] Clearly only a_1 matters. Had we defined the payoff by including $\|x^*x - (x^*x)^2)\|_{\mathrm{OP}}$ and $\|x^*x - xx^* - 1\|_{\mathrm{OP}}$ among the test-formulas of (P2), this would be a 'one move win' game for the Challenger. The payoff in an EF-game is sometimes defined by using a prescribed finite set of formulas; these games have numerous variations suitable for every occasion.

[11] Evaluation of continuous first-order statements results in real numbers instead of truth values; see [2], [5].

[12] There are better estimates but I am just a set theorist.

is that whether m has a fixed k as a divisor is determined by a single sentence, and the game associated with it has a fixed (not depending on m) number of innings and a fixed $\varepsilon > 0$.

How different is Γ_{HS} from Γ_{OP}? Can the Challenger distinguish the parities of the dimensions in this game? Well, no.

Lemma 2.3. *For all n and $\varepsilon > 0$ and all sufficiently large m, Duplicator wins $\Gamma_{HS}(\mathbb{M}_m, \mathbb{M}_{m+1}, n, \varepsilon)$.*

Proof. Let $\alpha_m \colon \mathbb{M}_m \to \mathbb{M}_{m+1}$ be defined by adding a zero $m+1$-st column and a zero $m+1$-st row to an $m \times m$-matrix and let $\beta_m \colon \mathbb{M}_{m+1} \to \mathbb{M}_m$ be defined by removing the $m+1$-st column and the $m+1$-st row from given $(m+1) \times (m+1)$ matrix. If $a \in \mathbb{M}_{m+1}$ and $\|a\|_{\mathrm{OP}} \leq 1$ we then have that $\|a - \alpha(\beta(a))\|_{\mathrm{HS}} < \sqrt{2/m}$, because the sum of the squares in the entries in each row and each column is at most 1. We can now describe a strategy for the Duplicator: In the j-th inning play $\alpha(a_j)$ or $\beta(b_j)$, depending on what the Challenger played. For a large enough m, this assures that after n moves the values in the payoff set hadn't changed by more than ε, resulting in a win for Duplicator. $\square$

This implies that the relations $v^*v = (v^*v)^2$ and $v^*v + vv^* = 1$ are (even when taken together) not weakly stable with respect to $\|\cdot\|_{\mathrm{HS}}$. A similar observation applies to divisibility of the dimension by any given $k \geq 3$. What, if any, properties of large matrix algebras can be detected by their $\|\cdot\|_{\mathrm{HS}}$-first order theories? Nobody knows.

Question 2.4. *Given n and $\varepsilon > 0$, is there $\bar{m}$ large enough so that Duplicator wins $\Gamma_{HS}(\mathbb{M}_l, \mathbb{M}_m, n, \varepsilon)$ whenever $\min(m, k) \geq \bar{m}$?*

A positive answer to this question would be equivalent to asserting that for every first-order sentence φ the limit $\lim_{m\to\infty} \varphi^{\mathbb{M}_m}$ exists.

3. Normalizers

> If you can't solve a problem, then there is an easier problem you can solve: find it.
>
> George Pólya

Let's change the game by requiring the players to play only permutation matrices. A *permutation matrix* is a matrix all of whose entries are equal to 0 or to 1 that has exactly one 1 in each row and each column. The linear transformation associated with such matrix has the effect of permuting the basis vectors. All permutation matrices are unitary and they are closed under products. Thus they form a multiplicative group in $\mathbb{M}_m$, and this group is isomorphic to $\mathcal{S}_m$, the group of permutations of an m-element set. On $\mathcal{S}_m$ we have the *Hamming metric* defined by (let $\Delta(\pi, \sigma)$ be the cardinality of the set $\{j : \pi(j) \neq \sigma(j)\}$)

$$d_H(\pi, \sigma) = \frac{\Delta(\pi, \sigma)}{m}.$$

Let $\Gamma(\mathcal{S}_m, \mathcal{S}_l, n, \varepsilon)$ be defined as follows.

- (N1) The players alternate choosing elements of $\mathcal{S}_m$ and $\mathcal{S}_l$, as before.
- (N2) Duplicator wins if and only if $|d_H(a_ia_j, a_k) - d_H(b_ib_j, b_k)| < \varepsilon$ for all i, j, and k.

This game is more similar to $\Gamma_{\mathrm{HS}}(\mathbb{M}_m, \mathbb{M}_l, n, \varepsilon)$ than it may appear. First, since we are comparing groups it is only natural to ignore the addition and multiplication by scalars. Second, the distance d_H is similar to the metric that one gets by identifying π and σ with permutation matrices and computing the HS-norm of the difference.[13] The latter is equal to $\sqrt{2\Delta(\pi,\sigma)/n}$. However, with fixed l, m, and n, for all small enough $\varepsilon > 0$ replacing (N2) with the following would not change the outcome of the game.

(N3) Duplicator wins if and only if $|\|a_i a_j - a_k\|_{\mathrm{HS}} - \|b_i b_j - b_k\|_{\mathrm{HS}}| < \varepsilon$ for all i, j, and k.

We can now answer the analog of Question 2.4.

Theorem 3.1. *There are n and $\varepsilon > 0$ such that there are arbitrarily large m and l for which the Challenger wins $\Gamma(\mathcal{S}_m, \mathcal{S}_l, n, \varepsilon)$.*

This is a consequence of [1, Theorem 5.5], although this is not at all obvious from what we have seen so far. The proof involves some fascinating ideas that we have no time to discuss such as sofic groups ([7]), expander graphs, Ulam-stability of approximate homomorphisms, comparison of first-order theories of alternating groups of order $p^4 - 1$ for primes $p \geq 7$, and something that awaits us in the final section. It is the time to take the red pill,[14] step out of the matrix, and take a bird's eye view.

4. The really, really, really, big picture

I have lied to you. Question 2.4 is just a shadow of the following question asked by Sorin Popa, and negative answer to Question 2.4 implies a negative answer to Question 4.1.[15]

Question 4.1. *Are all tracial ultraproducts of matrix algebras $\prod^{\mathcal{U}} \mathbb{M}_n$ isomorphic?*

It seems that I have a lot of explaining to do (pun intended). A nonprincipal ultrafilter on $\mathbb{N}$ (all ultrafilters in this snapshot are nonprincipal and on $\mathbb{N}$) is a family of subsets $\mathcal{U}$ of $\mathbb{N}$ that satisfies the following conditions.

($\mathcal{U}$1) For all A, B in $\mathcal{U}$, $A \cap B \in \mathcal{U}$.
($\mathcal{U}$2) For all $A \in \mathcal{U}$ and $A \subseteq C \subseteq \mathbb{N}$, $C \in \mathcal{U}$.
($\mathcal{U}$3) For every $A \subseteq \mathbb{N}$ exactly one of A or $\mathbb{N} \setminus A$ belongs to $\mathcal{U}$.
($\mathcal{U}$4) No finite set belongs to $\mathcal{U}$.

The first two conditions assert that $\mathcal{U}$ is a filter, while the third and fourth make it 'ultra' and nonprincipal, respectively. Such an object cannot be constructed without appealing to a fragment of the Axiom of Choice. While the latter has a number of fairly perverse consequences, the vast majority of contemporary mathematicians accept it as true. Every ultrafilter has remarkable properties. For example, if (x_n) is a bounded sequence of scalars, then there is a unique scalar x such that for every $\varepsilon > 0$ the set $\{n : |x - x_n| < \varepsilon\}$ belongs to $\mathcal{U}$. In this case we write $\lim_{n\to\mathcal{U}} x_n = x$ and say that the sequence (x_n) converges to x as $m \to \mathcal{U}$, or that it converges to x along $\mathcal{U}$. This readily implies that for every m

---

[13] Third, from the unitary group of an operator algebra such as $\mathbb{M}_m$ one can often recover the algebra, but that is another story.

[14] After reading this far you may have an impression that it was written by a fan of the Matrix movies. This is not true at all; I just had to grab your attention somehow. Sorry.

[15] Is the converse true? Who knows. Read on.

and $\|\cdot\|_{\mathrm{OP}}$-bounded sequence $a_n \in \mathbb{M}_m$ there is a unique $a \in \mathbb{M}_m$ such that $\lim_{n\to\mathcal{U}} \|a_m - a\|_{\mathrm{HS}} = \lim_{n\to\mathcal{U}} \|a_m - a\|_{\mathrm{OP}} = 0$. More generally, every sequence (x_n) in a compact metric (or any compact Hausdorff) space has a unique limit along $\mathcal{U}$, denoted $\lim_{n\to\mathcal{U}} x_n$.

'Limits' along an ultrafilter when compactness is lost are more interesting than limits in a compact space beyond comparison.

4.1. **Ultraproducts.** Given a sequence $G_n = (V_n, E_n)$, for $n \in \mathbb{N}$, of graphs, on the set $\prod_n V_n$ consider the relations

$$(a_n) \sim_{\mathcal{U}} (b_n) \Leftrightarrow \{n : a_n = b_n\} \in \mathcal{U},$$
$$(a_n)\, \mathrm{E}_{\mathcal{U}} (b_n) \Leftrightarrow \{n : a_n\, \mathrm{E}_n\, b_n\} \in \mathcal{U}.$$

Then $\sim_{\mathcal{U}}$ is an equivalence relation, and if $(a_n) \sim_{\mathcal{U}} (a'_n)$ and $(b_n) \sim_{\mathcal{U}} (b'_n)$ then $(a_n)\,\mathrm{E}_{\mathcal{U}}(b_n)$ if and only $(a'_n)\,\mathrm{E}_{\mathcal{U}}(b'_n)$. The quotient structure $(\prod_n V_n / \sim_{\mathcal{U}}, \mathrm{E}_{\mathcal{U}})$ is the *ultraproduct* $\prod_{\mathcal{U}} G_n$. For every first-order sentence φ about graphs, φ holds in $\prod_{\mathcal{U}} G_n$ if and only if the set $\{n : \varphi \text{ holds in } G_n\}$ belongs to $\mathcal{U}$. This is Łoś's Theorem, rightfully known as the Fundamental Theorem of Ultraproducts. Theorem 1.1 then implies that the first-order theory of $\prod_{\mathcal{U}} C_m$ does not depend on the choice of $\mathcal{U}$, the 0-1 law for random graphs implies the analogous statement for (sufficiently) random graphs, and by Theorem 1.2 every sequence of graphs has a further subsequence with this property. All of this applies to other discrete mathematical structures such as linear orderings, groups, rings..., but I owe you an explanation of Question 4.1.

Back on track, let[16]

$$\mathbb{M} = \{(a_m) : a_m \in \mathbb{M}_m \text{ and } \sup_m \|a_m\|_{\mathrm{OP}} < \infty\},$$
$$c_{\mathcal{U}}(\mathbb{M}) = \{(a_m) \in \textstyle\prod_m \mathbb{M}_m : \lim_{m\to\mathcal{U}} \|a\|_{\mathrm{HS}} = 0\}.$$

Both structures are closed under all algebraic operations defined coordinatewise, and $c_{\mathcal{U}}(\mathbb{M})$ is an *ideal* in $\mathbb{M}$: for $a \in \mathbb{M}$ and $b \in c_{\mathcal{U}}(\mathbb{M})$ we have $ab \in c_{\mathcal{U}}(\mathbb{M})$ and $ba \in c_{\mathcal{U}}(\mathbb{M})$. The *tracial ultraproduct* of the $\mathbb{M}_m$'s is the quotient

$$\prod\nolimits^{\mathcal{U}} \mathbb{M}_m = \mathbb{M}/c_{\mathcal{U}}(\mathbb{M}).$$

By the continuous analog of Łoś's Theorem, the value of every first-order sentence in the product is the limit of its values in $\mathbb{M}_m$ as $m \to \mathcal{U}$. It implies that the answer to Question 2.4 is positive if and only if the first-order theory of $\prod^{\mathcal{U}} \mathbb{M}_m$ does not depend on the choice of $\mathcal{U}$.

By [4], the negation of Cantor's Continuum Hypothesis (CH) implies a negative answer to Question 4.1. Instead of offering a detour on CH this late in the paper, I recommend that the next time you meet a set theorist you ask whether CH is true. You may hear answers that are about as opinionated as they are different.[17]

A remarkable property of all ultraproducts associated with nonprincipal utrafilters on $\mathbb{N}$ is their *countable saturation* (see [9] for an enlightening exposition). All that we need to know is a consequence of this property, that if $\prod^{\mathcal{U}} \mathbb{M}_n$ and $\prod^{\mathcal{V}} \mathbb{M}_n$

[16]The reason why we consider only operator norm-bounded sequences (and the reason for the similar requirement in (P1)) is that this assures the resulting algebra is isomorphic to an algebra of bounded operators on a Hilbert space. This algebra even happens to be a II_1 factor.

[17]Also see the Paul Erdös anecdote in [3, p. 238]. This book also happens to be my favourite reference for ultraproducts, countable saturation, and whatnot 😇.

share the first-order theory *and* the CH holds, then they are isomorphic.[18] Thus CH implies that Question 2.4 and Question 4.1 have the same answer. The answer to Question 2.4 does not depend on CH or other standard additional set-theoretic axioms; we just don't know what it is.

By replacing $\|\cdot\|_{\mathrm{HS}}$ with $\|\cdot\|_{\mathrm{OP}}$ in the definition of $c_{\mathcal{U}}(\mathbb{M})$, one obtains the (operator) norm ultraproduct of $\prod_{\mathcal{U}} \mathbb{M}_m$. Theorem 2.2 and the paragraph following it imply that for every $k \geq 2$ the first-order theory of $\prod_{\mathcal{U}} \mathbb{M}_m$ detects whether the set $\{m : k|m\}$ belongs to $\mathcal{U}$, and the operator norm analog of Question 4.1 has a negative answer regardless of whether CH is true or not. I don't know whether the theory of $\prod_{\mathcal{U}} \mathbb{M}_m$ can detect any other information about $\mathcal{U}$.

Acknowledgements. I am indebted to Dima Shlyakhtenko for suggesting that I write something like this and for helpful remarks on the first draft of this paper, to Diego Bejarano and Andrea Vaccaro for noticing a few typos, and to Vadim Alekseev for two insightful suggestions that improved the exposition.

References

[1] V. Alekseev and A. Thom. On non-isomorphic universal sofic groups. Preprint, arXiv:2406.06741 [math.GR] (2024), 2024.

[2] I. Ben Yaacov, A. Berenstein, C.W. Henson, and A. Usvyatsov. Model theory for metric structures. In Z. Chatzidakis et al., editors, *Model Theory with Applications to Algebra and Analysis, Vol. II*, number 350 in London Math. Soc. Lecture Notes Series, pages 315–427. London Math. Soc., 2008.

[3] I. Farah. *Combinatorial Set Theory and C*-algebras*. Springer Monographs in Mathematics. Springer, 2019.

[4] I. Farah, B. Hart, and D. Sherman. Model theory of operator algebras II: Model theory. *Israel J. Math.*, 201:477–505, 2014.

[5] B. Hart. An introduction to continuous model theory. In *Model theory of operator algebras*, pages 83–131. Berlin: De Gruyter, 2023.

[6] W. Hodges. *Model theory*, volume 42 of *Encyclopedia of Mathematics and its Applications*. Cambridge university press, 1993.

[7] V. Pestov. Hyperlinear and sofic groups: a brief guide. *Bull. Symbolic Logic*, 14:449–480, 2008.

[8] S. Shelah and J. Spencer. Zero-one laws for sparse random graphs. *J. Am. Math. Soc.*, 1(1):97–115, 1988.

[9] T. Tao. Ultraproducts as a bridge between hard analysis and soft analysis. https://terrytao.wordpress.com/tag/countable-saturation/, 2011.

Department of Mathematics and Statistics, York University, 4700 Keele Street, North York, Ontario, Canada, M3J 1P3, and Matematički Institut SANU, Kneza Mihaila 36, 11 000 Beograd, p.p. 367, Serbia

Email address: ifarah@yorku.ca

URL: https://ifarah.mathstats.yorku.ca

[18]This is a part of the evidence that the first-order logic lies in the 'Goldilocks expressiveness zone'. The nonisomorphic ultraproducts obtained when CH fails can be distinguished only by considering very large and subtle invariants coming from Saharon Shelah's non-structure theory.