



Zero-One Laws for Sparse Random Graphs

Author(s): Saharon Shelah and Joel Spencer

Source: *Journal of the American Mathematical Society*, Vol. 1, No. 1 (Jan., 1988), pp. 97-115

Published by: [American Mathematical Society](#)

Stable URL: <http://www.jstor.org/stable/1990968>

Accessed: 17/06/2014 05:01

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



American Mathematical Society is collaborating with JSTOR to digitize, preserve and extend access to *Journal of the American Mathematical Society*.

<http://www.jstor.org>

ZERO-ONE LAWS FOR SPARSE RANDOM GRAPHS

SAHARON SHELAH AND JOEL SPENCER

1. INTRODUCTION AND SUMMARY

Let $G = G(n, p)$ denote the random graph with n vertices and edge probability p . Let GRA denote the first order theory of graphs, as described at the end of this section. For any $p \in [0, 1]$ and any statement A in GRA set

$$f(n, p, A) = \Pr[G(n, p) \text{ has } A].$$

Fagin [6] and, independently, Glebskii et al. [7] proved that for all p, A ,

$$\lim_{n \rightarrow \infty} f(n, p, A) = 0 \text{ or } 1.$$

For random graph theorists (see, e.g., Bollobás [1] for general reference) p “any constant” is not the only, not even the most interesting case. Rather, they consider $p = p(n)$, a function approaching zero. In their seminal paper, Erdős and Rényi [5] showed that for many interesting A there is a function $p(n)$, which they called a threshold function, so that if $r(n) \ll p(n)$ then $f(n, r(n), A) \rightarrow 0$ while if $p(n) \ll r(n)$ then $f(n, r(n), A) \rightarrow 1$. (Notation: $p \ll r$ means $\lim p/r = 0$. All limits are as n approaches infinity.) Let us say $p = p(n)$ satisfies the Zero-One Law if for all A in GRA, $\lim f(n, p, A) = 0$ or 1 . We shall partially characterize those $p = p(n)$ which satisfy the Zero-One Law.

When $p < n^{-1+o(1)}$, the “very sparse” range, we give in §2 a nearly complete answer.

Theorem 1. *If either*

- (i) $p \ll n^{-2}$,
- (ii) for some integer k , $n^{-1-1/k} \ll p \ll n^{-1-1/(k-1)}$,
- (iii) $n^{-1-\varepsilon} \ll p \ll n^{-1}$ for all $\varepsilon > 0$,

Received by the editors January 15, 1987.

1980 *Mathematics Subject Classification* (1985 Revision). Primary 05C80, 03C10.

Research of first author partially supported by U. S.-Israel Binational Science Foundation.

Research of second author partially supported by National Science Foundation (U.S.).

©1988 American Mathematical Society
0894-0347/88 \$1.00 + \$.25 per page

- (iv) $n^{-1} \ll p \ll n^{-1} \log n$, or
- (v) $n^{-1} \log n \ll p \ll n^{-1+\varepsilon}$ for all $\varepsilon > 0$,

then p satisfies the Zero-One Law.

The functions n^{-2} , $n^{-1-1/k}$, n^{-1} , and $n^{-1} \log n$ are all known threshold functions. In this range the Zero-One Law holds as long as p falls “between the cracks” of the spectrum of threshold functions. For completeness we mention that Fagin’s proof actually gives that if $p \gg n^{-\varepsilon}$ for all $\varepsilon > 0$ then p satisfies the Zero-One Law.

It had been conjectured by the second author that if $p(n)$ were reasonably smooth then $\lim f(n, p, A)$ would exist. In §3 we give the following counterexample to that conjecture.

Theorem 2. *There is a statement A so that for any $p = n^{-1/7} q^{1/7}$ with $n^{-1/\log_5 n} < q < \log n / \log_5 n$ the function $f(n, p, A)$ does not approach a limit in n . Moreover, if we restrict n to satisfy $\log^* n \equiv 25 \pmod{100}$ then $f(n, p, A) \rightarrow 1$ whereas if we restrict n to satisfy $\log^* n \equiv 75 \pmod{100}$ then $f(n, p, A) \rightarrow 0$.*

Here $\log_5 n = \log \log \log \log \log n$ and $\log^* n$ is the least k so that the sequence $u_0 = n$, $u_{i+1} = \log u_i$ has $u_k < 2$ where, for convenience, we assume all logs are to base two.

The statement A will be (theoretically) explicitly given; it allows us to essentially interpret a segment of arithmetic on an appropriate intersection of neighborhoods. The exponent $-1/7$ and the expression $\log_5 n$ are conveniences; for any rational r , $0 < r < 1$, we may find an A so that for p appropriately near n^{-r} $f(n, p, A)$ does not approach a limit in n .

In §4 we prove

Theorem 6. *If α is irrational then $p = n^{-\alpha}$ satisfies the Zero-One Law.*

In the final section we take a dynamic view and examine the behavior of $\text{Pr}[A]$ as p evolves from 0 to 1 through the functions of n . We define a spectrum $\text{Spec}(A)$, giving those spots where $\text{Pr}[A]$ changes, and give a partial characterization of the possible spectra of first order statements.

The first order theory of graphs, GRA, consists of two binary predicates, equality and adjacency, with adjacency assumed symmetric and antireflexive. The formulae of GRA are built up from atomic formulae of type either $x_i = x_j$ or $x_i \sim x_j$, where \sim denotes adjacency, using the logical connectives \wedge (and), \vee (or), \neg (not), \rightarrow (implies) and also the quantifiers \exists (there is), \forall (for all). Here the x_i are interpreted as the vertices of the underlying graph (“the universe”) so quantifiers may be read “there is a vertex x_i such that” and “for every vertex x_i there holds.” Formulae in which all occurring variables are bounded by quantifiers are called closed formulae or sentences. One method, the method we employ, to prove results on formulae is by induction along the

way it is built up. By propositional calculus every formula is equivalent to another built just by using \neg , \vee , and \exists .

GRA is our framework throughout this paper. In GRA we can make statements such as “ G contains a triangle” or “ G has no isolated points.” Other statements, such as “ G is planar” or “ G is Hamiltonian,” cannot be made in this language. This restriction may appear artificial to many graph theorists who feel that appropriate Zero-One Laws hold for a much wider class of “natural” graph-theoretic statements. There are logical pitfalls. Kaufmann and Shelah [9] have shown that even Fagin’s Theorem does not hold in monadic second order logic. It would be interesting to find a logical structure stronger than GRA in which, for example, the statements above could be made and for which analogs to Fagin’s Theorem and our results could be proven.

2. VERY SPARSE GRAPHS

Here we prove Theorem 1, which is really five theorems. The plan, due to Fagin, is identical in all cases. If we can find a family \mathcal{A} of statements in GRA such that

- (i) $\lim f(n, p, A) = 1$ for all $A \in \mathcal{A}$,
- (ii) \mathcal{A} is complete, i.e., every statement of GRA is either provable or refutable from \mathcal{A} ,

then p satisfies the Zero-One Law. For let B be any statements in GRA. Suppose B is provable from \mathcal{A} . By compactness (the finiteness of proof) some finite number A_1, \dots, A_m of statements in \mathcal{A} will imply B . As

$$f(n, p, A_1 \wedge \dots \wedge A_m) \geq 1 - \sum_{i=1}^m f(n, p, \neg A_i),$$

$\lim f(n, p, A_1 \wedge \dots \wedge A_m) = 1$ and hence $\lim f(n, p, B) = 1$. If B is not provable then by completeness $\neg B$ is provable, $\lim f(n, p, \neg B) = 1$ and $\lim f(n, p, B) = 0$.

In each case we give a suitable \mathcal{A} . Verification of (i) is then an elementary exercise in random graphs. To show (ii) it suffices, by the Gödel Completeness Theorem, to show

- (iii) all countable models of \mathcal{A} are isomorphic.

Fagin let \mathcal{A} be the set of extension axioms $A_{r,s}$: For all distinct $x_1, \dots, x_r, y_1, \dots, y_s$ there is a point z adjacent to all of the x_i and none of the y_j . The countable G satisfying all $A_{r,s}$ are isomorphic by a simple “back and forth” argument.

Case 1. $p \ll n^{-2}$. \mathcal{A} consists of the single statement “There are no edges.”

Case 2. $n^{-1-1/k} \ll p \ll n^{-1-1/(k-1)}$. For each tree T on at most k points (including the tree on one vertex) and each r let $A_{T,r}$ be the statement “ G has (at least) r components T .” Let B be the statement “ G does not contain $k+1$ points whose internal edges contain a tree.” Let C be the statement “ G

has no cycles of size at most k ." Let \mathcal{A} be all of the above statements. A countable model G must then consist precisely of a countable number of copies of every tree T of size at most k and (iii) is satisfied.

In the remaining cases we cannot show (iii) but rather

(iv) all countable models of \mathcal{A} are elementarily equivalent.

Proof of (iv) in the cases below can be given in various ways, see, e.g., Marcus [8], particularly Lemma 2.1.

Case 3. $n^{-1-\varepsilon} \ll p \ll n^{-1}$ for all $\varepsilon > 0$. For each tree T and r let $A_{T,r}$ be the statement "There are (at least) r components T ." For each r let C_r be the statement "There are no cycles of length r ." A countable model G then consists of countably many copies of each finite tree plus possibly some infinite trees.

Case 4. $n^{-1} \ll p \ll n^{-1} \log n$. For each finite tree T let $A_{T,r}$ be the statement "There are (at least) r components T ." For all $m \geq 3$, r let $B_{m,r}$ be the statement "There are (at least) r cycles of size m ." For all m let C_m be the statement "There do not exist m vertices containing (at least) $m+1$ internal edges." For all m, r let $D_{m,r}$ be the statement "There do not exist m vertices with exactly m internal edges with one vertex having (total) degree precisely r ." A countable model G consists of countably many copies of each finite tree, for each m countably many components consisting of a unique cycle of length m with all vertices having infinite degree plus possibly some infinite trees.

Remark. N. Pippenger (San Jose) has noted that the case $p = \Theta(n^{-1} \ln n)$ contains a doubly infinite sequence of "tight" threshold functions. Let \mathcal{F} be the class of functions

$$f(n) = \frac{1}{K} \frac{\ln n}{n} + \frac{L}{K} \frac{\ln \ln n}{n}$$

where either $K = 1$, $L = 0$ or $K \geq 2$, $L \geq K - 1$, both integral. Say $p(n) <^* f(n)$ if $\lim n(f(n) - p(n)) = +\infty$. The Zero-One Law holds in this range for $p(n)$ if for every $f(n) \in \mathcal{F}$ either $p(n) <^* f(n)$ or $f(n) <^* p(n)$.

Case 5. $n^{-1} \log n \ll p \ll n^{-1+\varepsilon}$ for all $\varepsilon > 0$. For each r let A_r be the statement "There are no points of degree r ." Let B_r be the statement "No r points have $r+1$ internal edges." For each $m \geq 3$, r let $C_{r,m}$ be the statement "There exist r cycles of length m ." Let G_m be that unique connected graph containing a single cycle of length m with all points of infinite degree. Let G_0 be the tree with all points of infinite degree. A countable model G consists of countably many copies of each G_m and possibly some copies of G_0 .

3. INTERPRETING ARITHMETIC

The proof of Theorem 2 proceeds in three stages. First we define A . Second we show that the random graph almost always has certain properties. Finally,

we show that when G has these properties A holds when $\log^* n \equiv 25 \pmod{100}$ and A does not hold when $\log^* n \equiv 75 \pmod{100}$.

Statement A will have the form

$$\exists_{x_1, \dots, x_7, y, y_1, \dots, y_6} \text{STRAT} \wedge \text{ARITH} \wedge \text{MAX} \wedge \text{LOGSTARCOND}.$$

We use a number of auxiliary predicates to define A , some of which depend on x_1, \dots, y_6 . All variables are distinct unless otherwise stated.

$$S(z_1, \dots, z_7): \neg(\exists w)(w \sim z_1 \wedge \dots \wedge w \sim z_7),$$

$$N_1 = \{z: x_1 \sim z \wedge \dots \wedge x_7 \sim z\}.$$

Membership in N_1 is expressible in GRA. We use y, y_1, \dots, y_6 for the (somewhat technical) task of stratifying N_1 .

STRAT: For all $z_1 \in N_1$ there exist unique $z_2, \dots, z_6 \in N_1$ with $S(y, z_1, \dots, z_6)$. Furthermore, for all $z_1 \in N_1$, there is a unique i , $1 \leq i \leq 6$ for which there exist z_2, \dots, z_6 with $S(y_i, z_1, \dots, z_6)$. (We shall say z_1 is associated with i in that case.) Furthermore, if $S(y, z_1, \dots, z_6)$ then z_1, \dots, z_6 are associated with distinct i .

Set N equal to the set of $z \in N_1$ associated with $i = 1$. For $z \in N$ set $z^{(1)} = z$ and $z^{(i)}$ equal to that w associated with i so that $S(y, z, w, z_3, \dots, z_6)$ for some $z_3, \dots, z_6 \in N_1$. On N define the *ordered* 6-ary predicate (here the z_i may be equal)

$$H(z_1, \dots, z_6): S(x, z_1^{(1)}, z_2^{(2)}, \dots, z_6^{(6)}).$$

ARITH: There is a labelling $1, 2, \dots, s$ of N so that $H(z_1, \dots, z_6)$ if and only if $z_1 + z_2 = z_3$ and $z_4 z_5 = z_6$.

Expression of ARITH in GRA is an elementary exercise in logic. We want to say that N is the largest such set over all x_1, \dots, y_6 . If membership in N, N^* is expressible then we write

BIGGER (N, N^*): There exist v_3, \dots, v_7 so that, writing $S(x, y)$ for $S(x, y, v_3, \dots, v_7)$ the relation S gives an injection from $N - N^*$ to $N^* - N$ which is not a bijection.

In an interpretation with N, N^* finite sets if **BIGGER** (N, N^*) then N^* is indeed bigger (has more elements) than N , though the converse may not hold. Now define

MAX: There do not exist x_1^*, \dots, y_6^* satisfying **STRAT** and **ARITH** so that, with N^* defined analogously to N , **BIGGER** (N, N^*).

Let $\text{tower}(x)$ be the number-theoretic function defined inductively by $\text{tower}(1) = 2$ and $\text{tower}(x+1) = 2^{\text{tower}(x)}$ so that $\text{tower}(x)$ is an exponential tower of twos of height x . On N we develop arithmetic via the Godel β -function and let **TOWER**(x, y) be the predicate with interpretation $y = \text{tower}(x)$. Define a unary predicate

$$\text{PSEUDOLOGSTAR}(x): (\exists y)\text{TOWER}(x, y) \wedge \neg(\exists y)\text{TOWER}(x+1, y).$$

Finally, still quantifying over N

$$\text{LOGSTARCOND: PSEUDOLOGSTAR}(x) \rightarrow (\exists q) \\ (x = 100q + 1 \vee x = 100q + 2 \vee \cdots \vee x = 100q + 50)).$$

This defines the statement A .

We now proceed to properties of random graphs. For any G set

$$N(y) = \{z : \{y, z\} \in G\}, \quad N_Y = \bigcap_{y \in Y} N(y)$$

and let H^7 be the 7-graph

$$H^7 = \{Y \in [n]^7 : N_Y = 0\}.$$

For every vertex x let H_x denote the 6-graph

$$H_x = \{Y \in [n]^6 : Y \cup \{x\} \in H\}.$$

For any $T \subset [n]$ with $x \notin T$ let $H_x(T)$ denote the restriction of H_x to T .

Let us say G has property **UNIV** (for universal) if for every $T \subset [n]$ with $|T| \leq \log_8 n$ (recall special notation of §1) and every 6-graph H on T there exists $x \notin T$ so that $H_x(T) = H$. Let us say G has property **SIZE** if for every $t \leq \log_8 n$ there is a 7-set X with $|N_X| = t$. All probabilities will refer to the random graph $G(n, p)$ with p satisfying $p = n^{-1/7} q^{1/7}$ and

$$n^{-1/\log_5 n} < q < \log n / \log_5 n.$$

The choice of $\log_8 n$ above is not “best possible” but simply a function that grows much slower than $\log_5 n$. (The case $q = 1$ already contains the basic argument.)

Claim.

$$\Pr[\text{UNIV}] = 1 - o(1), \quad \Pr[\text{SIZE}] = 1 - o(1).$$

We outline the arguments which use only standard random graph methods. For T, H, x as above let $A(T, H, x)$ be the event $H_x(t) = H$ and let $A(T, H)$ be the event that for no x , $A(T, H, x)$ holds. Fix T, H . Our object, toward the first claim, is to bound $\Pr[A(T, H)]$ from above. Call G **NICE** if

$$\frac{1}{2}np^6 < |N_X| < 2np^6 \quad \text{for all } X \in [n]^6, \\ |N_X| < 10np^7 \log n \leq \log^3 n \quad \text{for all } X \in [n]^7.$$

For fixed X , $|N_X|$ has binomial distribution $B(n - 6, p^6)$ and $B(n - 7, p^7)$ respectively. The classic Chernoff [3] bounds on the tail of the binomial distribution give that the probability of the above conditions failing goes rapidly to zero, even when multiplied by the t^6 or t^7 possible X . We use only

$$\Pr[\text{not NICE}] = o(n^{-\log_7 n})$$

and we bound

$$\Pr[A(T, H)] = o(n^{-\log_7 n}) + \Pr[A(t, H) | G \text{ NICE}].$$

We shall show

$$(*) \quad \Pr[A(T, H)] = o(n^{-\log_7 n}).$$

It suffices to show $(*)$ conditional on G NICE. We further condition on any particular values for $N(x)$, $x \in T$. Let R denote the union of all N_X over all $X \in [T]^7$ so that, by NICEness, R has at most $\binom{t}{7} \log^3 n < \log^4 n$ elements. For $X \in [T]^6$ let $N'_X = N_X - R$ so that all N'_X have between $0.4np^6$ and $2np^6$ elements.

Let $A^*(T, H, x)$ be the event

- (i) $N(x) \cap R = \emptyset$.
- (ii) If $X \subset T$, $|X| = 6$, and $X \in H$ then $N(x) \cap N'_X = \emptyset$.
- (iii) If $X \subset T$, $|X| = 6$, and $X \notin H$ then $N(x) \cap N'_X \neq \emptyset$.

Since $|R| < \log^4 n$, $\Pr[N(x) \cap R = \emptyset] > 1 - p|R| = 1 - o(1)$. The bounds on $|N'_X|$ give

$$\begin{aligned} \Pr[N(x) \cap N'_X = \emptyset] &= (1 - p)^{|N'_X|} > e^{-3np^7} = e^{-3q}, \\ \Pr[N(x) \cap N'_X \neq \emptyset] &= 1 - (1 - p)^{|N'_X|} > 1 - e^{-0.3np^7} > 0.2q. \end{aligned}$$

Here we see the asymmetry between $q > 1$ and $q < 1$. Set

$$\alpha = \min(e^{-3q}, 0.2q).$$

As R and the N'_X are mutually disjoint the above events are mutually independent and

$$\Pr[A^*(T, H, x)] > (1 - o(1))\alpha^{\binom{t}{6}} > \alpha^{t^6}.$$

If $A^*(T, H, x)$ holds then $A(T, H, x)$ holds so $\Pr[A(T, H, x)] > \alpha^{t^6}$. The events $A(T, H, x)$ are mutually independent over all $x \notin T$, $x \notin N_X$ for all $X \subset T$, $|X| = 6$. There are at least $n/2$ such x so

$$\Pr[A(T, H)] < (1 - \varepsilon)^{n/2} \quad \text{with } \varepsilon = \alpha^{t^6}.$$

Now $(*)$ follows immediately from the bounds on q . There are fewer than $n^{\log_8 n}$ possible T, H . The probability that UNIV does not hold is at most $n^{\log_8 n} n^{-\log_7 n} = o(1)$, giving the first claim.

An overview of the above argument may be useful. Essentially H^7 behaves as a random 6-graph with edge probability $(1 - p^7)^n \sim e^{-q}$. With $e^{-q} < \frac{1}{2}$, $\Pr[H_x = H] \geq (e^{-q})^{\binom{t}{6}}$, the minimum achieved when $H = \emptyset$. With $e^{-q} > \frac{1}{2}$, $\Pr[H_x = H] \geq (1 - e^{-q})^{\binom{t}{6}}$, the minimum achieved when H is complete. With q near one $\alpha = \min[e^{-q}, 1 - e^{-q}]$ is not too small so $\Pr[H_x = H] \geq \varepsilon = \alpha^{\binom{t}{6}}$ is not too small. There are nearly n potential x and the H_x are independent so UNIV fails with probability at most $(1 - \varepsilon)^n$, which is very small.

SIZE is easier. Fix $n^{-0.1}$ disjoint sets X_i of size 7 and let A_i be the event that N_{X_i} has precisely t elements. Let B be the event that there are no edges on $\bigcup X_i$ so

$$\Pr[B] = (1 - p)^{\binom{7n^{0.1}}{2}} = o(1).$$

Conditioning on B , $|N_{X_i}|$ has binomial distribution $B(n - 7n^{0.1}, p^7)$ and so

$$\begin{aligned} \Pr[A_i|B] &= \binom{n - 7n^{0.1}}{t} (p^7)^t (1 - p^7)^{n - 7n^{0.1} - t} \\ &\sim (q^t/t!) e^{-q} > n^{-0.01} \end{aligned}$$

for q in this range. The A_i are then mutually independent so

$$\Pr[\bigwedge \bar{A}_i | B] < (1 - n^{-0.01})^{n^{0.1}} \ll 1$$

and

$$\Pr[\bigwedge \bar{A}_i] \leq \Pr[\bar{B}] + \Pr[\bigwedge \bar{A}_i | B] \ll 1.$$

Quite easily, $\Pr[\bigwedge \bar{A}_i] = o(1/\log_5 n)$. As there are only $\log_8 n$ values of T to obtain to realize **SIZE**, $\Pr[\neg \text{SIZE}] = o(1)$.

Finally, we come to the interpretation. Let G be any graph on n vertices satisfying **UNIV** and **SIZE**. We claim

- (i) If $\log^* n \equiv 25 \pmod{100}$ then G satisfies A .
- (ii) If $\log^* n \equiv 75 \pmod{100}$ then G does not satisfy A .

As almost all G have properties **UNIV** and **SIZE** this will conclude the proof of Theorem 2.

For (i), set $t = 6[(\log_8 n)/3]$. By **SIZE** there exist x_1, \dots, x_6 so that N_{X_i} has size t . Label N_{X_i} by u^i , $1 \leq i \leq 6$, $1 \leq u \leq t/6$ arbitrarily. Let H be the 6-graph on N_X consisting of the sets $\{u^1, u^2, u^3, u^4, u^5, u^6\}$. By **UNIV** there is a vertex $y \notin N_X$ so that $H_y(N_X)$ is precisely H^i . Let H^+ be the 6-graph of all $\{a^1, b^2, c^3, d^4, e^5, f^6\}$ with $a + b = c$ and $de = f$. Again by **UNIV** there is an x such that H_x is precisely H^+ . These x_1, \dots, y_6 satisfy **STRAT** and **ARITH**. Maybe **MAX** is not satisfied. However, now take among all choices of x_1, \dots, y_6 for which **STRAT** and **ARITH** are satisfied that one for which N has the biggest size. As some choice gives $|N| = \lfloor \log_8 n/3 \rfloor$ the maximal has $|N| = s \geq \lfloor \log_8 n/3 \rfloor$. This choice will satisfy **MAX** since $BIGGER(N, N^*)$ implies that N^* is really bigger than N . On this N , if $TOWER(x, y)$ then $y = \text{tower}(x)$ so $x \leq \log^* s$. The converse may not hold since $TOWER(x, y)$ requires numbers larger than y in using the Gödel β -function. However, it is easy to show that no numbers larger than 2^y are needed. That is: if $\text{tower}(x) = y$ and $2^y < s$ then $TOWER(x, y)$. Thus $PSEUDOLOGSTAR(x)$ implies $\log^* s - 1 \leq x \leq \log^* s$. As $\log_9 n \equiv 25 \pmod{100}$, x is between 15 and 25 mod 100 and A holds.

For (ii), we know there exist x_1^*, \dots, y_6^* so that **STRAT** and **ARITH** are satisfied with $N_X^* = \lfloor \log_8 n/3 \rfloor$. Suppose x_1, \dots, y_6 satisfy **STRAT** and **ARITH**

and give a set N_X with $|N_X| < [\log_8 n/3]$. Pick $v_3, \dots, v_6 \in [n] - N_X - N_X^*$ arbitrarily and let $T = N_X \cup N_X^* \cup \{v_3, \dots, v_6\}$ so that $|T| \leq \log_8 n$. As N_X^* is bigger than N_X there is a function f from $N_X - N_X^*$ into $N_X^* - N_X$ which is an injection but not a surjection. Let H be the 6-graph on T consisting of all sets $\{x, f(x), v_3, \dots, v_6\}$, $x \in N_X - N_X^*$. By **UNIV** there is a $v_7 \notin T$ so that $H = H_{v_7}$. Hence **BIGGER**(N_X, N_X^*) is satisfied. (**BIGGER** is bigger on small sets!) Then **MAX** is false. That is, any x_1, \dots, y_6 such that **STRAT** and **ARITH** and **MAX** are satisfied has N_X with s elements, $s \geq [\log_8 n/3]$. Hence that element x with **PSEUDOLOGSTAR**(x) satisfies $\log^* n - 10 \leq x \leq \log^* n$. With $\log^* n \equiv 75 \pmod{100}$, x must lie between 65 and 75 mod 100; hence A is false.

The analysis of $\lim f(n, p, A)$ for this A may be extended to $p = n^{-1/7-\varepsilon}$, ε a small constant. With notation as before define properties

UNIV $_\varepsilon$: For every $T \subset [n]$ with $|T| \leq \varepsilon^{-1/10}$ and every 6-graph H on T there exists $x \notin T$ so that $H_x(T) = H$.

SIZE $_\varepsilon$: For every $t < \varepsilon^{-1/2}$ there is a 7-set X with $|N_X| = t$.

A.a. both **UNIV** $_\varepsilon$ and **SIZE** $_\varepsilon$ hold. (The exponents of ε are not best possible.) Furthermore, a.a. $|N_1| \leq \varepsilon^{-1}$ for all x_1, \dots, x_7 . If x_1, \dots, y_6 are such that N satisfies **STRAT** and **ARITH** and **MAX** then $\varepsilon^{-1/11} < |N| < \varepsilon^{-1}$. It follows that:

- (i) If $[\log^* \varepsilon^{-1}] \equiv 25 \pmod{100}$ then $\lim f(n, p, A) = 1$.
- (ii) If $[\log^* \varepsilon^{-1}] \equiv 75 \pmod{100}$ then $\lim f(n, p, A) = 0$.

4. IRRATIONAL EXPONENT

In this section we prove that $p = n^{-\alpha}$, α irrational, satisfies the Zero-One Law. Throughout this section α is a fixed irrational, $0 < \alpha < 1$, $p = n^{-\alpha}$ and G is the random graph $G(n, p)$. An event occurs almost always—abbreviated a.a.—if its probability approaches unity in n . We let $H_1, E(H_1)$ denote the vertex and edge sets of H_1 . A rooted graph is a pair (H_0, H_1) where H_1 is a graph and H_0 is a subset of H_1 . (H_0, H_1) has type (v, e) if $v = |H_1 - H_0|$ and e is the number of $\{x, y\} \in E(H_1)$ with $y \notin H_0$. (Basically, the edges inside H_0 are to be disregarded.) We employ the interval notation: $[H_0, H_1]$ is the set of H , $H_0 \subset H \subset H_1$; $(H_0, H_1], [H_0, H_1)$, (H_0, H_1) are the same except H_0, H_1 , and both H_0 and H_1 respectively are to be excluded from the interval.

Definitions. (H_0, H_1) is *dense* if it has type (v, e) and $e/v > \alpha^{-1}$.

(H_0, H_1) is *sparse* if it has type (v, e) and $e/v < \alpha^{-1}$.

(H_0, H_1) is *safe* if (H_0, H) is sparse for all $H \in (H_0, H_1]$.

(H_0, H_1) is *rigid* if (H, H_1) is dense for all $H \in [H_0, H_1)$.

(H_0, H_1) is *hinged* if it is safe but for all $H \in (H_0, H_1)$ (H, H_1) is not safe.

We now give some elementary consequences of these definitions. When $H \in (H_0, H_1)$ and (H_0, H) , (H, H_0) , (H_0, H_1) have types (v', e') , (v'', e'') , (v, e)

respectively, counting gives

$$(**) \quad v = v' + v'', \quad e = e' + e''.$$

(A1) If (H_0, H) and (H, H_1) are sparse then (H_0, H_1) is sparse.

(A2) If (H_0, H) and (H, H_1) are dense then (H_0, H_1) is dense.

Proofs. Immediate from (**).

(B1) If (H_0, H_1) is sparse then (H, H_1) is safe for some $H \in [H_0, H_1]$.

Proof. Take $H \subset H_1$ maximal with (H_0, H) dense. (If no such H exists take $H = H_0$.) If (H, H_1) is not safe some (H, H') is dense but (A2) gives (H_0, H') dense, contradicting maximality.

(B2) If (H_0, H_1) is dense then (H_0, H) is rigid for some $H \in (H_0, H_1]$.

Proof. Take $H \supset H_0$ minimal with (H, H_1) sparse. (If no such H exists take $H = H_1$.) If (H_0, H) is not rigid some (H', H) is sparse but (A1) gives (H', H_1) sparse, contradicting maximality.

(C) If (H_0, H_1) is hinged of type (v, e) , $H \in (H_0, H_1)$, and (H_0, H) has type (v', e') then $e'/v' \leq e/v$.

Proof. If not, $e'/v' > e/v$ and (H, H_1) as type $(e - e')/(v - v') < e/v$ and so is sparse. (B1) gives $H^* \in [H, H_1]$ with (H^*, H_1) safe, contradicting definition of hinged.

(D) If (H_0, H_1) is dense and $X \cap H_1 = \emptyset$ then $(H_0 \cup X, H_1 \cup X)$ is dense.

Proof. Let (H_0, H_1) have type (v, e) . Then $(H_0 \cup X, H_1 \cup X)$ has type (v, e') where, as every edge counted in (H_0, H_1) is still counted in $(H_0 \cup X, H_1 \cup X)$, $e \leq e'$.

(E) If (H_0, H) and (H, H_1) are rigid then (H_0, H_1) is rigid.

Proof. Let $H^* \in [H_0, H_1]$. As (H_0, H_1) is rigid $(H \cap H^*, H)$ is dense so $(H^*, H \cup H^*)$ is dense, applying (D) with $X = H^* - H$. As (H, H_1) is rigid, $(H \cup H^*, H_1)$ is dense. By (A2), (H^*, H) is dense.

(F) If (H_0, H_1) is hinged and $H \in (H_0, H_1)$ then (H, H_1) is rigid.

Proof. If not, (H', H_1) is sparse for some $H' \in [H, H_1]$. By (B), (H'', H_1) is safe for some $H'' \in [H, H_1]$, contradicting the definition of hinged.

(G) If (H_0, H_1) is rigid and $X \cap H_1 = \emptyset$ then $(H_0 \cup X, H_1 \cup X)$ is rigid.

Proof. Let $H^* \in (H_0 \cup X, H_1 \cup X]$. As (H_0, H_1) is rigid, $(H^* - X, H_1)$ is dense, so, by (D), $(H^*, H_1 \cup X)$ is dense.

(H) Let Ω be an arbitrary graph, H_0 a subset of the vertex set of Ω . Let $H_i \subset (H_0, \Omega]$, (H_0, H_i) rigid, $1 \leq i \leq s$. Then $(H_0, H_1 \cup \dots \cup H_s)$ is rigid.

Proof. Assume $s = 2$. As (H_0, H_2) is rigid, $(H_1 \cap H_2, H_2)$ is rigid so, applying (G) with $X = H_1 - H_2$, $(H_1, H_1 \cup H_2)$ is rigid. As (H_0, H_1) is rigid, (E) implies $(H_0, H_1 \cup H_2)$ is rigid.

For $s > 2$ we use induction.

Given (H_0, H_1) we define the l -closure $\text{cl}_l(H_0; H_1)$ as the union of all sets $H \in (H_0, H_1]$ with (H_0, H) rigid and $|H - H_0| \leq l$. When there is no such H we define $\text{cl}_l(H_0; H_1) = H_0$. We often write $\text{cl}_l(H_0)$ when there is an understood universal set H_1 .

(I) $(H_0, \text{cl}_l(H_0))$ is rigid.

Proof. Immediate from (H).

(J) Suppose $H_0 \subset H_1$, $y \in H_1 - H_0$ and $(H_0 \cup \{y\}, H_1)$ is rigid. Set $H = \text{cl}_{|H_1 - H_0|}(H_0; H_1)$. Assume $y \notin H$. Then (H, H_1) is safe.

Proof. If not (H, H') is dense for some $H' \in (H, H_1]$. By (B2) (H, H'') is rigid for some $H'' \in (H, H']$. By (I) (H_0, H) is rigid so by (E) (H_0, H'') is rigid which, by definition of closure, would force $H'' \subset H$, a contradiction.

Definition. Let (H_0, H_1) be a rooted graph and $f: H_0 \rightarrow G$ an injection. We say $g: H_1 \rightarrow G$ is a *set-extension* if

(i) $x \in H_0 \Rightarrow g(x) = f(x)$.

We say g is an *extension* if, in addition,

(ii) $\{x, y\} \in E(H_1)$, $y \notin H_0 \Rightarrow \{g(x), g(y)\} \in E(G)$.

We say g is contained in B (equivalently, B contains g) if $g(H_1 - H_0) \subset B$.

Remark. For technical reasons we do *not* require f to map edges into edges.

Definition. $N(f, H_0, H_1)$ is the number of extensions g .

Theorem 3. Let (H_0, H_1) be safe. There exist c_1, c so that a.a. for all $f: H_0 \rightarrow G$

$$n^{v-\alpha e} (\log n)^{-c} < N(f, H_0, H_1) < c_1 n^{v-\alpha e}.$$

Proof. If (H_0, H_1) is not hinged let $H \in (H_0, H_1)$ so that (H, H_1) is safe. As (H_0, H_1) is safe, (H_0, H_1) is safe a fortiori. Each $f: H_0 \rightarrow G$ can be extended to $g: H \rightarrow G$ in $N(f, H_0, H)$ ways and thence to $h: H_1 \rightarrow G$ in $N(g, H, H_1)$ ways so the result follows by induction on $|H_1 - H_0|$. Henceforth, we assume (H_0, H_1) hinged.

We first show the lower bound. A pair (f, B) with $f: H_0 \rightarrow G$ and $B \subset G$ is called *bad* if there is no extension $g: H_1 \rightarrow G$ contained in B .

Lemma 1. *There is a constant $k > 0$ with the following property. Set $m = kn^{ae/v}$ and let $f: H_0 \rightarrow G$, $B \subset G$ with $|B| = m$ be fixed. Then*

$$\Pr[(f, B) \text{ bad}] < \frac{1}{2}.$$

Proof. For convenience delete $f(H_0)$ from B , adjusting k if necessary, so that $f(H_0) \cap B = \emptyset$.

For each $S \subset B$ with size $|H_1 - H_0|$ specify arbitrarily a set-extension g with $g(H_1 - H_0) = S$. Let $X_S = 1$ if g_S is an extension; otherwise zero. Let $X = \sum X_S$. Then $E[X_S] = p^e$ and

$$E(X) = \binom{m}{v} p^e > k^v / v!.$$

We want $\Pr[X = 0] < \frac{1}{2}$ for which we calculate the second moment.

$$\begin{aligned} \text{Var}(X) &= \sum_S \text{Var}(X_S) + \sum_{S \neq T} \text{Cov}(X_S, X_T) \\ &\leq E(X) + \sum_{j=0}^{v-1} \sum_{|S \cap T|=j} \text{Cov}(X_S, X_T). \end{aligned}$$

When $j = 0$ X_S, X_T are independent and $\text{Cov}(X_S, X_T) = 0$. Otherwise

$$\text{Cov}(X_S, X_T) \leq E(X_S X_T) = p^u.$$

Here u is the number of pairs which are images of $\{x, y\} \in E(H_1)$, $y \notin H_0$ under g_S or g_T . As g_S, g_T are bijections there are e such pairs in each case so $u = 2e - u'$ where u' is the number of pairs in both categories. Set $H = g_S^{-1}(S \cap T)$ so that $|H| = j$. Then u' counts at most edges $\{x, y\} \in E[H \cup H_0]$ with $y \notin H_0$. By (C)—using here that (H_0, H_1) is hinged— $u'/j \leq e/v$.

$$\sum_{|S \cap T|=j} \text{Cov}(X_S, X_T) \leq m^{2v-j} p^{2e-j(e/v)} = [m^v p^e]^{2-j/v}$$

so

$$\text{Var}(X) \leq k^v / v! + \sum_{j=1}^{v-1} (k^v)^{2-j/v} < \frac{1}{2} E(X)^2,$$

when k is chosen sufficiently large. By Chebyshev's Inequality

$$\Pr[(f, B) \text{ bad}] \leq \Pr[X = 0] \leq \text{Var}(X) / E(X)^2 \leq \frac{1}{2}.$$

Lemma 2. *Fix $f: H_0 \rightarrow G$, $C \subset G$ with $|C| = k' n^{ae/v} \log n$. For k' sufficiently large a.a.*

$$\Pr[(f, C) \text{ bad}] > \frac{1}{2} n^{-|H_0|}.$$

Proof. Split C into $k'' \log n$ disjoint B_i of size m given by Lemma 1. For (f, C) to be bad all (f, B_i) must be bad. But the events “ (f, B_i) bad” involving disjoint edge sets are mutually independent, so

$$\Pr[(f, C) \text{ bad}] < 2^{-k'' \log n},$$

which can be made smaller than any power of n .

With k' as above there are only $n^{|H_0|}$ injections $f: H_0 \rightarrow G$ so a.a. for every f at most half the (f, C) are bad. Set $M = |C| = k' n^{\alpha e/v} \ln n$ for convenience. Each extension $g: H_1 \rightarrow G$ is contained in $\binom{n-v}{M-v} \sim \binom{n}{m} (M/n)^v$ m -sets. Together at most $|N(f, H_0, H_1)| \binom{n-v}{M-v}$ M -sets contain an extension but at least $\frac{1}{2} \binom{n}{M}$ do. Thus

$$|N(f, H_0, H_1)| \left(\frac{M}{n}\right)^v \geq \frac{1}{2},$$

$$|N(f, H_0, H_1)| \geq \frac{1}{2} \left(\frac{n}{M}\right)^v > n^{v-\alpha e} (\ln n)^{-v+o(1)},$$

completing the proof of the lower bound of Theorem 3.

Conjecture. A.a. for all f $|N(f, H_0, H_1)| > c' n^{v-\alpha e}$ where c' depends only on H_0 and H_1 .

The upper bound to Theorem 3 also requires some preliminary lemmas. Two extensions g_1, g_2 of f on (H_0, H_1) are called *disjoint* if

$$g_1(H_1 - H_0) \cap g_2(H_1 - H_0) = \emptyset.$$

Lemma 3. Let (H_0, H_1) be dense. There exist K so that a.a. for all $f: H_0 \rightarrow G$ there do not exist K disjoint extensions g_1, \dots, g_K .

Proof. Let $c = |H_0|$ and (H_0, H_1) have type (v, e) . Pick K so that $c + K(v - \alpha e) < 0$. There are less than $n^c (n^v)^K (f, g_1, \dots, g_K)$ with g_i disjoint set-extensions of f . Each g_i is an extension with probability p^e and since they are disjoint they are all extensions with probability p^{eK} so the probability of (f, g_1, \dots, g_K) existing with all g_i extensions is less than $n^{c+vK-\alpha eK} = o(1)$.

Lemma 4. Let (H_0, H_1) be rigid. There exist K so that a.a. for all $f: H_0 \rightarrow G$ there do not exist K extensions g_1, \dots, g_K .

Proof. For all $H \in (H_0, H_1)$, (H, H_1) is dense so we may pick K_0 so that Lemma 3 holds for all (H, H_1) . By Erdős and Rado [4] we may pick K so that given any K functions on $v = |H_1 - H_0|$ points there is a Δ -system of K_0 functions. These K_0 functions would be disjoint extensions of the common root H .

Now we fix $f: H_0 \rightarrow G$ and bound $s' = |N(f, H_0, H_1)|$ from above. For any $h \in H_1$, $y \in G$ $(H_0 \cup \{h\}, H_1)$ is rigid (F) so by Lemma 4 there are at most K extensions g with $g(h) = y$. Each extension g is disjoint from all but at most $K |H_0|^2$ extensions and so there is a set of $s = s'/(K|H_0|^2 + 1)$ disjoint extensions.

There are less than $n^c \binom{n}{v}^s / s!$ pairs consisting of an $f: H_0 \rightarrow G$ and a set $\{g_1, \dots, g_s\}$ of disjoint set-extensions of f . The probability that all g_i are extensions is, as before, p^{es} . Thus the probability of $f, \{g_1, \dots, g_s\}$ existing is at most

$$n^c n^{vs} p^{es} / s! < n^c (2.72 n^v p^{e/s})^s \ll 1$$

when $s = 3n^{v-\alpha e}$. Thus, a.a.,

$$N(f, H_0, H_1) \leq 3(K|H_0|^2 + 1)n^{v-\alpha e},$$

completing the proof of Theorem 3.

Lemma 5. *For all r, l there is a K so that a.a. $|\text{cl}_l(G_0)| \leq K$ for all $G_0 \subset G$, $|G_0| \leq r$.*

Proof. There are a bounded (Lemma 4) number of extensions for each rigid (H_0, H_1) with $|H_0| = |G_0|$ and only a finite number of such (H_0, H_1) up to isomorphism.

Lemma 6. *For all r, l_1, l_2 there is an l so that a.a.*

$$\text{cl}_{l_1}(\text{cl}_{l_2}(G_0)) \subseteq \text{cl}_l(G_0)$$

for all $G_0 \subset G$, $|G_0| \leq r$.

Proof. Set $l = K + l_1$ where $|\text{cl}_{l_2}(G_0)| \leq K$ by Lemma 5. If $y \in \text{cl}_{l_1}(\text{cl}_{l_2}(G_0))$ then $y \in G'$, $|G' - \text{cl}_{l_2}(G_0)| \leq l_1$, $(\text{cl}_{l_2}(G_0), G')$ rigid. From (I) and (E), (G_0, G') is rigid, $|G' - G_0| \leq l$ so $y \in \text{cl}_l(G_0)$.

Let $x_1, \dots, x_r \in G$. By $\text{cl}_l(x_1, \dots, x_r)$ we shall mean the graph on $\text{cl}_l(\{x_1, \dots, x_r\})$ with points x_1, \dots, x_r specified. We say (x_1, \dots, x_r) , (x'_1, \dots, x'_r) have the same l -type if $\text{cl}_l(x_1, \dots, x_r) \cong \text{cl}_l(x'_1, \dots, x'_r)$, the isomorphism sending x_i to x'_i . By Lemma 5 the family $\widehat{L}_{r,l}$ of possible l -types is a.a. bounded. Each one can be considered as a graph H_1 with specified vertices h_1, \dots, h_r such that $H_1 = \text{cl}_l(\{h_1, \dots, h_r\}; H_1)$.

Now we are ready to deal with first order sentences. Let $\phi(x_1, \dots, x_r)$, as customary, denote a formula in GRA with free variables x_1, \dots, x_r .

Theorem 4. *For every $\phi(x_1, \dots, x_r)$ there is an l so that a.a. for all $x_1, \dots, x_r \in G$ $\phi(x_1, \dots, x_r)$ depends only on $\text{cl}_l(x_1, \dots, x_r)$. That is, there exist l , $\mathcal{A} \subseteq \widehat{L}_{r,l}$ so that a.a.*

$$(***) \quad G \models \phi(x_1, \dots, x_r) \Leftrightarrow \text{cl}_l(x_1, \dots, x_r) \in \mathcal{A}.$$

Example. Let $\frac{1}{3} + \frac{1}{9}k < \alpha < \frac{1}{3} + \frac{1}{9}(k-1)$ and set

$$\begin{aligned} \phi(x, y, z): (\exists w)(\forall a)[a \sim x \wedge a \sim y \wedge a \sim z \rightarrow (\exists! b \neq a)(\exists v) \\ b \sim x \wedge b \sim y \wedge b \sim z \wedge v \sim a \wedge v \sim b \wedge v \sim w], \end{aligned}$$

$\phi(x, y, z)$ depends on $\text{cl}_{3k+1}(x, y, z)$. $\text{cl}_l(x, y, z)$ contains all points a_1, \dots, a_{2s} adjacent to x, y, z . (If there are an odd number of such points $G \models \neg \phi$.) If $s < k$ then $G \models \phi$ since the extension by w, v_1, \dots, v_s with $w \sim v_i \sim a_{2i-1}, a_{2i}$ is safe. When $s \geq k$, $G \models \phi$ if and only if $\text{cl}_{3k+1}(x, y, z)$ contains points w, v_1, \dots, v_s so that, after renumbering, $w \sim v_i \sim a_{2i-1}, a_{2i}$. An examination

of $\text{cl}_{3k}(x, y, z)$ would not suffice since if $\text{cl}_{3k}(x, y, z)$ consists of a_1, \dots, a_{2k} joined to x, y, z the desired w, v_1, \dots, v_k might exist outside $\text{cl}_{3k}(x, y, z)$.

Proof. We use induction on the length of ϕ . Let us say l, \mathcal{A} give ϕ if (***) holds. Let $\neg, \vee, (\exists y)$ be the only logical connectives. For atomic formulae $x_i \sim x_j$ or $x_i = x_j$ simply take $l = 0$. If $\phi = \neg\psi$ and l, \mathcal{A} give ψ then $l, \widehat{L} - \mathcal{A}$ give ϕ . If $l < l'$ then $\text{cl}_{l'}(x_1, \dots, x_r)$ determines $\text{cl}_l(x_1, \dots, x_r)$ so that if l, \mathcal{A} give ϕ and $l < l'$ there is an \mathcal{A}' so that l', \mathcal{A}' give ϕ . If $\phi = \phi_1 \vee \phi_2$ and, by induction, l_1, \mathcal{A}_1 give ϕ_1 and l_2, \mathcal{A}_2 give ϕ_2 then set $l = \max(l_1, l_2)$ so that l, \mathcal{A}_1' give ϕ_1 and l, \mathcal{A}_2' give ϕ_2 . ϕ is then given by $l, \mathcal{A}_1' \cup \mathcal{A}_2'$. This leaves the critical case $\phi(x_1, \dots, x_r): (\exists y)\psi(x_1, \dots, x_r, y)$. By the inductive hypothesis there exist $l, \mathcal{A} \subseteq \widehat{L}_{r+1, l}$ giving ψ . Thus

$$G \models \phi(x_1, \dots, x_r) \Leftrightarrow \text{for some } y \in G \text{ type}_l(x_1, \dots, x_r, y) \in \mathcal{A}.$$

It suffices to show the following.

Theorem 5. Let $H_0 = (h_1, \dots, h_r, h^*)$ and let $(H_0, H_1) \in \widehat{L}_{r+1, l}$. Then there exists l' and $\mathcal{B} \subseteq \widehat{L}_{l', r}$ so that a.a. for all $x_1, \dots, x_r \in G$

$$(\exists y)\text{type}_l(x_1, \dots, x_r, y) = (H_0, H_1) \Leftrightarrow \text{type}_{l'}(x_1, \dots, x_r) \in \mathcal{B}.$$

Assume Theorem 5 and let \mathcal{A} , above, consist of types T^1, \dots, T^s . For each T^i there are l^i, \mathcal{B}^i satisfying Theorem 5. We replace all l^i by $l = \max(l^1, \dots, l^s)$ and \mathcal{B}^i by \mathcal{B}^{i*} . Then ϕ is given by l, \mathcal{A} where $\mathcal{A} = \mathcal{B}^{1*} \cup \dots \cup \mathcal{B}^{s*}$.

To prove Theorem 5 we set

$$H = \text{cl}_{|H_1|}[\{h_1, \dots, h_r\}; H_1].$$

Case 1. $h^* \in H$. As $(\{h_1, \dots, h_r, h^*\}, H_1)$ is rigid by (E) $(\{h_1, \dots, h_r\}, H_1)$ is also rigid and so $H = H_1$. Let l' be such that $\text{cl}_l(\text{cl}_{l'}(X)) \subset \text{cl}_{l'}(X)$ for all $X \subset G$, $|X| = r$. Let \mathcal{B} be the set of all l' -types $((h_1, \dots, h_r), H_2)$ so that for some $h^* \in H_2$

$$((h_1, \dots, h_r, h^*, \text{cl}_l(\{h_1, \dots, h_r, h^*\}); H_2)) \cong (H_0, H_1).$$

If $\text{type}_l(x_1, \dots, x_r) \in \mathcal{B}$ then with y the image of h^* $\text{type}_l(x_1, \dots, x_r, y) = (H_0, H_1)$. Conversely if there exists y with $\text{type}_l(x_1, \dots, x_r, y) = (H_0, H_1)$ then $y \in \text{cl}_l(x_1, \dots, x_r)$ so $\text{cl}_l(x_1, \dots, x_r, y) \subset \text{cl}_{l'}(x_1, \dots, x_r)$ and so $\text{cl}_{l'}(x_1, \dots, x_r) \in \mathcal{B}$.

Case 2. $h^* \notin H$. We claim

$$\begin{aligned} (\exists y)\text{type}_l(x_1, \dots, x_r, y) &= ((h_1, \dots, h_r, h^*), H_1) \\ &\Leftrightarrow \text{type}_l(x_1, \dots, x_r) = ((h_1, \dots, h_r); H). \end{aligned}$$

As $\text{type}_l(x_1, \dots, x_r, y)$ determines $\text{type}_l(x_1, \dots, x_r)$, \Rightarrow is immediate. Now assume $\text{type}_l(x_1, \dots, x_r) = ((h_1, \dots, h_r), H)$; we must find a “witness” y . By (J) the rooted graph (H, H_1) is safe. Fix $f: H \rightarrow G$ with $f(h_i) = x_i$ and $\{h', h''\} \in E(H) \Leftrightarrow \{f(h'), f(h'')\} \in E(G)$. (f exists by definition of type.) Let (H, H_1) have type (v, e) . Theorem 3 bounds the number of extensions $g: H_1 \rightarrow G$. If $g: H_1 \rightarrow G$ is an extension, setting $y = g(h^*)$ might not work as $\text{cl}_l(x_1, \dots, x_r, y)$ might have more vertices or edges than $(H_1, E(H_1))$. This is critical—we need to find a witness y with the desired special properties and no others.

Let \mathcal{H} be the family of types $((h_1, \dots, h_r, h^*), H^*) \in \widehat{L}_{r+1, l}$ such that H_1 is a subgraph of H^* . Let $M(H^*)$ be the number of y so that there is an extension $g: H^* \rightarrow G$ of f relative to (H_0, H^*) with $g(h^*) = y$. For each y Lemma 4 gives there are at most K such f . Let (v^*, e^*) denote the type of (H_0, H^*) . Theorem 3 gives

$$n^{v^* - \alpha e^*} (\ln n)^{-c} < M(H^*) < c_1 n^{v^* - \alpha e^*}$$

for adjusted constants c, c_1 . There are $M(H_1)$ y such that $\text{cl}_l(x_1, \dots, x_r, y)$ contains $(H_1, E(H_1))$ as a subgraph. Let \mathcal{H}^* be \mathcal{H} with $(H_1, E(H_1))$ deleted. For each $(H^*, E(H^*)) \in \mathcal{H}^*$ there are at most $M(H^*)$ y so that $\text{cl}_l(x_1, \dots, x_r, y) \cong (H^*, E(H^*))$. Thus there are at least

$$M(H) - \sum_{H^* \in \mathcal{H}^*} M(H^*)$$

y with type $\text{cl}_l(x_1, \dots, x_r, y) = ((h_1, \dots, h_r), H_1)$.

Let $(H^*, E(H^*)) \in \mathcal{H}^*$ have type (v^*, e^*) . Suppose $H_1 \subsetneq H^*$. Since $(\{h_1, \dots, h_r, h^*\}, H^*)$ is rigid so is (H_1, H^*) , hence $(e^* - e)/(v - v^*) > \alpha^{-1}$ and $v^* - e^* \alpha < v - \alpha e$. When $H_1 = H^*$ we have $v = v^*$ so $e^* > e$ and again $v^* - e^* \alpha < v - \alpha e$. In all cases $M(H^*)$ is much smaller than $M(H)$. Then

$$\begin{aligned} M(H) - \sum_{H^* \in \mathcal{H}^*} M(H^*) &> n^{v - \alpha e} (\ln n)^{-c} - \sum c_1 n^{v^* - \alpha e^*} \\ &= n^{v - \alpha e + o(1)} > 0. \end{aligned}$$

The witness y exists, giving Case 2, Theorem 5 and Theorem 4.

Theorem 6. If α is irrational then $p = n^{-\alpha}$ satisfies the Zero-One Law.

Proof. Let ϕ be a closed formula. By Theorem 4 there exist l, \mathcal{A} so that a.a.

$$G \models \phi \Leftrightarrow \text{cl}_l(\emptyset) \in \mathcal{A}.$$

A dense rooted graph (\emptyset, H_1) of type (v, e) appears an expected number $n^v p^e = o(1)$ of times in G . For a given l there are only a fixed number of such H_1 so a.a. no such H_1 appears. That is, a.a. $\text{cl}_l(\emptyset) = \emptyset$. Hence a.a.

$$G \models \phi \Leftrightarrow \emptyset \in \mathcal{A}.$$

Either $\emptyset \in \mathcal{A}$ or $\emptyset \notin \mathcal{A}$ so either a.a. $G \models \phi$ or a.a. $G \models \neg\phi$.

5. THRESHOLD SPECTRA

The analysis of §4 was essentially static. Now consider A fixed and consider the property A over the “evolution” of the random graph. We would like to describe this evolution in terms of a spectrum of threshold functions, though the results of §3 force us to use some care.

Call $a > 0$ a point of continuity if there is an $\varepsilon > 0$ and $\delta \in \{0, 1\}$ so that if $n^{-a-\varepsilon} < p < n^{-a+\varepsilon}$ then $\text{Lim } f(n, p, A) = \delta$. Otherwise, call a an evolutionary discontinuity and define the *threshold spectrum* $\text{Spec}(A)$ as the set of such a . From the methods of §4 it follows that all irrational a are points of continuity and that $\text{Spec}(A)$ is a closed nowhere dense set of rational numbers. Many interesting A analyzed by Erdős, Rényi, and others have threshold spectrum consisting of a single point. For monotone A , not necessarily in GRA, the existence of a threshold function is shown by Bollobás and Thomason [2] to follow from purely combinatorial argument. The analysis concluding §3 shows that $\text{Spec}(A)$ may have a limit point.

A set of real numbers S is called *scattered* if it is closed and the sequence $S = S_0, S_1, \dots$, with S_{i+1} the set of limit points of S_i , has $S_m = \emptyset$ for some m .

Theorem 7. $\text{Spec}(A)$ is scattered.

We outline the argument, which requires proving Theorem 6 in a manner uniform on α . The example following the statement of Theorem 4 illustrates that the l of Theorem 4 need not be bounded uniformly in α . The essential induction of Theorem 4 becomes the following.

Lemma 7. For every $\phi(x_1, \dots, x_r)$ there is a scattered $S = \text{Scat}(\phi) \subset Q$ so that for all a, b with $[a, b] \cap S = \emptyset$ there is an l, \mathcal{A} so that if $n^{-b} < p < n^{-a}$ then a.a. for all $x_1, \dots, x_r \in G$

$$G \models \phi(x_1, \dots, x_r) \Leftrightarrow \text{cl}_l(x_1, \dots, x_r) \in \mathcal{A}.$$

The proof is by induction. We let $S = \{0, 2\}$ for the null statement to set up boundaries. The cases $\phi = \neg\psi$ and $\phi = \phi_1 \vee \phi_2$ are simple; again the critical case is

$$\phi(x): (\exists y)\psi(x, y),$$

where we abbreviate x_1, \dots, x_r as x . Let $S_0 = \text{Scat}(\psi)$. We expand S_0 in four stages.

For each consecutive $a, a' \in S_0$ add a doubly infinite monotonic sequence b_n , $n \in \mathbb{Z}$, of rationals from a to a' . To be explicit: $b_n = a' - (a' - a)2^{-n-1}$ for $n \geq 0$; $b_n = a + (a' - a)2^{n-1}$ for $n < 0$. Let S_1 be the enlarged set. For every consecutive $b, b' \in S_1$ (note: the $a \in S_0$ are now limit points so $b = b_n$,

$b' = b_{n+1}$) find, by induction, l, \mathcal{A} so that for $n^{-b'} < p < n^{-b}$ a.a. for all x

$$(\text{****}) \quad G \models \phi(x) \Leftrightarrow (\exists y) \text{cl}_l(x, y) \in \mathcal{A}.$$

When $(H_0, H_1) \in \mathcal{A}$, $H_0 = (h_1, \dots, h_r, h^*)$ we consider

$$H = \text{cl}_{|H_1|}[\{h_1, \dots, h_r\}; H_1]$$

as a function of $\alpha \in [b', b]$. Fix M with $M \geq r + l$ and $M \geq |H_1|$ for all $(H_0, H_1) \in \mathcal{A}$. In $[b', b]$ add all rationals x/y with $y \leq \binom{M}{2}$, giving S_2 . The value of H can change only at $\alpha \in S_2$. For consecutive $c, c' \in S_2$ (****) still holds and H is well defined over $\alpha \in (c, c')$. When $h^* \notin H$

$$(\exists y) \text{cl}_l(x, y) = (H_0, H_1) \Leftrightarrow \text{cl}_l(x) = ((h_1, \dots, h_r); H).$$

For each consecutive $c, c' \in S_2$ add a doubly infinite sequence d_n of rationals from c to c' giving S_3 . Let d, d' be consecutive elements of S_3 . All rooted graphs (H, H_1) with $|H_1| \leq r + l$ with type (v, e) that are dense over $\alpha \in (c, c')$ will have $v - \alpha e < -\varepsilon$ where ε is constant over all $\alpha \in [d, d']$. Lemma 3 gives a K so that no $f: H \rightarrow G$ has more than K disjoint extensions to $g: H_1 \rightarrow G$. Lemmas 4, 5, and 6 follow and there is an l^* so that if $n^{-d'} < p < n^{-d}$ then a.a. $\text{cl}_l(\text{cl}_l(x)) \subset \text{cl}_{l^*}(X)$ for all $X \subset G$, $|X| = r$. When $h^* \in H$ the statement $(\exists y) \text{cl}_l(x, y) = (H_0, H_1)$ is determined by $\text{cl}_{l^*}(x)$. For every consecutive d, d' add all rational $x/y \in (d, d')$ with $y \leq \binom{r+l^*}{2}$, giving S_4 .

Assume $[A, B] \cap S_4 = \emptyset$. There is an $\varepsilon > 0$ so that $|\alpha - x/y| > \varepsilon$ for all x, y above. As before, there is a constant K so that if $n^{-B} < p < n^{-A}$ then a.a. $|\text{cl}_{l^*}(X)| \leq K$ for all $X \subset G$, $|X| = r$. Then

$$(\exists y) \text{cl}_l(x, y) = (H_0, H_1) \Leftrightarrow \text{cl}_{l^*}(x) \in \mathcal{B}$$

where \mathcal{B} is that set of possible closures that have an h^* generating H_1 .

As \mathcal{A} is a finite set and we can do this for all $(H_0, H_1) \in \mathcal{A}$, there is an l^*, \mathcal{B}^* so that with $n^{-B} < p < n^{-A}$ a.a.

$$(\exists y) \text{cl}_l(x, y) \in \mathcal{A} \Leftrightarrow \text{cl}_{l^*}(x) \in \mathcal{B}^*$$

and so we may set $\text{Scat}(\phi) = S_4$.

S is precisely the set of accumulation points of S_4 so S_4 is scattered by induction, completing Lemma 7.

Now let A be a closed sentence of GRA and let $S = \text{Scat}(A)$ be given by Lemma 7. Suppose $\alpha \notin S$. As S is closed some $[\alpha - \varepsilon, \alpha + \varepsilon] \cap S = \emptyset$. Hence there are l, \mathcal{A} so that for all p with $n^{-\alpha-\varepsilon} < p < n^{-\alpha+\varepsilon}$ a.a.

$$G \models A \Leftrightarrow \text{cl}_l(\emptyset) \in \mathcal{A}.$$

But, as before, a.a. $\text{cl}_l(\emptyset) = \emptyset$. If $\emptyset \in \mathcal{A}$ then a.a. $G \models A$, otherwise a.a. $G \models \neg A$. Hence $\alpha \notin \text{Spec}(A)$. $\text{Spec}(A)$ is a closed subset of a scattered set, hence it is scattered.

A full characterization of the threshold spectra of first order statements has proven elusive. In particular, we do not know if $\text{Spec}(A)$ may contain an infinite number of limit points.

BIBLIOGRAPHY

1. B. Bollobás, *Random graphs*, Academic Press, New York, 1985.
2. B. Bollobás and A. G. Thomason, *Threshold functions*, *Combinatorica* **7** (1986), 35–38.
3. H. Chernoff, *A measure of asymptotic efficiency for tests of a hypothesis based on a sum of observations*, *Ann. Math. Stat.* **23** (1952), 493–509.
4. P. Erdős and R. Rado, *Intersection theorems for systems of sets*, *J. London Math. Soc.* **35** (1960), 85–90.
5. P. Erdős and A. Rényi, *On the evolution of random graphs*, *Publ. Math. Inst. Hungar. Acad. Sci.* **5** (1960), 17–61.
6. R. Fagin, *Probabilities on finite models*, *J. Symbolic Logic* **41** (1976), 50–58.
7. Y. V. Glebskii, D. I. Kogan, M. I. Liogonkii and V. A. Talanov, *Range and degree of realizability of formulas in the restricted predicate calculus*, *Cybernetics* **5**, 142–154.
8. I. Marcus, *Minimal models of one function symbol*, *Israel J. Math.* **18** (1974), 117–130.
9. M. Kaufman and S. Shelah, *On random models of finite powers and monadic logic*, *Discrete Math.* **54** (1985), 285–293.

DEPARTMENT OF MATHEMATICS, HEBREW UNIVERSITY, JERUSALEM, ISRAEL

DEPARTMENT OF MATHEMATICS, SUNY AT STONY BROOK, STONY BROOK, NEW YORK 11794

COURANT INSTITUTE, 251 MERCER, NEW YORK, NEW YORK 10012 (Current address of Joel Spencer)