Ups and Downs of First Order Sentences on Random Graphs

JOEL SPENCER* GÁBOR TARDOS†

AMS Subject Classification: 05C80, 60F20
Keywords: random graph, first order logic, zero-one laws, polynomial time hierarchy

Abstract

Shelah and Spencer [1] proved that the zero-one law holds for the first order sentences on random graphs $G(n, n^{-\alpha})$ whenever α is a fixed positive *irrational.* This raises the question what zero-one valued functions on the positive irrationals arise as the limit probability of a first order sentence on these graphs. Here we prove two necessary conditions on these functions, a number-theoretic and a complexity condition. We hope to prove in a subsequent paper that these conditions together with two simpler and previously proved conditions are also sufficient and thus they constitute a characterization.

1 Introduction, results

In this paper we consider the limit probabilities of first order sentences on random graphs. Recall that the variables of the first order statements on a simple undirected graph G range over the vertices of G. The statements are built from the atomic formulae $x = y$ or $x \sim y$ (the latter interpreted as "x is adjacent to y") using logical connectives $\wedge$, $\vee$, $\neg$ and the quantifiers $\exists x$ and $\forall x$.

Consider the random graph $G(n, p)$ on n vertices with each edge present independently with probability p. Let us set $p = n^{-\alpha}$, α a fixed positive

*Courant Institute of Mathematical Sciences, 251 Mercer St., New York, NY 10012 USA. Email: spencer@cs.nyu.edu

†Mathematical Institute of the Hungarian Academy of Science, Reáltanoda u. 13-15, H-1055 Budapest, Hungary; Email: tardos@cs.elte.hu

irrational real. Let A be any first order sentence. Saharon Shelah and the first author of the present paper showed [1] that $\lim_{n\to\infty}\Pr[G(n,p)\models A]$ always exists and is equal to zero or one. Here we take a more evolutionary view, fixing A and considering how this limit varies as α changes. Formally, we let I denote the positive irrationals throughout and define $f_A : I \to \{0,1\}$ by

$$f_A(\alpha) = \lim_{n\to\infty} \Pr[G(n,n^{-\alpha}) \models A] \tag{1}$$

In this work we give a sequence of conditions that all functions $f = f_A$ must satisfy. We further believe that these conditions are sufficient—that for any function f satisfying the conditions given below there is a first order sentence A with $f = f_A$. That is, we believe that we have a *complete characterization* of the functions f_A. We hope to return to this in a second part of this paper.

In [3] the almost sure theory T_α for $G(n,n^{-\alpha})$ was given a combinatorial axiomatization (cf. Lemma 10, and the axioms preceeding it). For $\alpha \in I$ we have $f_\alpha(A) = 1$ if and only if $A \in T_\alpha$. We employ this axiomatization with the ironic consequence that our current work, outside of motivating comments and examples, is devoid of probabilistic calculation.

Examples. $A =$ "G contains a triangle". It is well known that n^{-1} is the threshold function for A so that $f_A(\alpha) = 0$ for $\alpha > 1$ and $f_A(\alpha) = 1$ for $\alpha < 1$. Note the value *at* $\alpha = 1$ is not considered. Let B be the property $\omega(G) = 3$. As $n^{-2/3}$ is the threshold function for containing a 4-clique $f_B(\alpha)$ is zero from ∞ to 1, then one from 1 to 2/3 and zero from 2/3 to 0. Observe that with the parameterization $p = n^{-\alpha}$ the evolution of the random graph goes "backwards". We may think of α starting at 2 (for $\alpha > 2$ there are no edges) and getting smaller. From $\alpha = 2$ to $\alpha = 1$ is the very sparse region described below. As α gets smaller and smaller (but still positive) the graph gets denser and denser. Close to zero we have the very dense region.

In this section we give the conditions on f_A without proof. Two regions of the domain are particularly simple.

`Very Sparse Condition:` f_A is constant on each interval $(1+\frac{1}{i+1}, 1+\frac{1}{i})$ and on $(2,\infty)$. Further there exists $k = k(A)$ such that f_A is constant on $(1, 1+\frac{1}{k})$

`Very Dense Condition:` There exists a positive integer $k = k(A)$ such that f_A is constant on $(0, \frac{1}{k})$.

Of course, the statement that f_A is constant on an interval really means it is constant on the irrationals of the interval, as that is where it is defined. We naturally say that f_A is continuous at a positive c if there is an $\epsilon > 0$ such that f_A is constant on $(c-\epsilon, c+\epsilon)$. This is defined for both rational and irrational c. As it was proved in [1] f_A is continuous at each irrational c. When f_A is not continuous at c we call c a point of discontinuity. With this notion we can rephrase the Very Sparse Condition as saying there are only a finite number of points of discontinuity ("ups and downs") in $(1, \infty)$ and they are all of the form $1 + \frac{1}{i}$ with i a positive integer. The Very Dense Condition can be expressed by saying that zero is not an accumulation point of the points of discontinuity.

Our core interest will be in the behavior of f_A on $(\frac{1}{k}, 1)$. Here the situation is considerably more complicated and more interesting. As in the Very Sparse situation the points of discontinuity are all rational numbers. But unlike the Very Sparse situation it is possible for f_A to have an infinite number of points of discontinuity. To describe the behavior we make a detour into elementary number theory that seems intriguing in its own right.

For $0 \leq \beta \leq \alpha$ real (though our interest is only in $0 < \alpha \leq 1$) we define

$$\tau(\alpha, \beta) = \sup\{\frac{k}{l} \leq \alpha \mid \frac{k-1}{l} \leq \beta\} \tag{2}$$

where the $k \geq 0$, $l > 0$ must be integers. We clearly have $\tau(\alpha, \alpha) = \alpha$. As we shall see (Lemma 2) for $\beta < \alpha$ the value $\tau(\alpha, \beta)$ is a rational from the interval $(\beta, \alpha]$. We consider $\tau(\alpha, \beta)$ the next rational approximation of the real α following β. We define an approximation sequence for each $\alpha \geq 0$ by setting $\tau_0(\alpha) = 0$ and inductively defining

$$\tau_{i+1}(\alpha) = \tau(\alpha, \tau_i(\alpha)). \tag{3}$$

This defines an infinite sequence of increasing rational approximations to α if α is irrational, but for rational α the sequence stabilizes in a finite number of steps when α is reached (Lemma 3). We define the first index i for which $\tau_i(\alpha) = \alpha$ as the *length* of the rational α.

`Discontinuity Condition:` There exists a positive integer $k = k(A)$ such that all points of discontinuity of f_A are rationals of length at most k.

We devote Section 2 to some observations related to the sequence $\tau_i(\alpha)$ and our notion of length of rational numbers. We introduce there several alternative notions of length and study if they are equivalent with respect to the Discontinuity Condition.

Let us define $LEN(k)$ to be those $\alpha \geq 0$ of length at most k. As we shall see in Lemma 5, $LEN(k)$ is closed and well-ordered under $>$. Thus for $\alpha \in LEN(k)$ we can define α^- as the maximal element of $LEN(k)$ strictly smaller than α. (Note that α^- depends on k. We use $\nu_k(\alpha)$ for α^- in Section 2.)

We can now rephrase the Discontinuity Condition as follows: There exists a positive integer $k = k(A)$ such that for all $\alpha \in LEN(k)$ the function f_A is constant on (α^-, α).

To prepare for our last condition for $f : I \to \{0,1\}$ and $\alpha > 0$ we define

$$f^-(\alpha) = \lim_{\epsilon \to 0^+} f(\alpha - \epsilon) \tag{4}$$

That is, if $f^-(\alpha)$ exists, then it is the constant $\delta \in \{0,1\}$ so that $f = \delta$ in the interval $(\alpha - \epsilon, \alpha)$ for ϵ appropriately small.

From the Discontinuity Condition f_A^- is well defined for all $\alpha > 0$. When $\alpha \in LEN(k)$ (with the value k as in the Discontinuity Condition) the value $f_A^-(\alpha)$ is the constant value of f_A on (α^-, α). As these intervals partition $(0,1)$ the function f_A on $(0,1]$ is determined by the function f_A^- on $LEN(k)$. Not all such functions are viable—for one thing there are only countably many A and $LEN(k)$ is infinite. Our final condition limits the appropriately defined complexity of f_A^-.

Consider f_A^- as a function from the rational numbers $\alpha > 0$ to $\{0,1\}$. Regard each $\alpha = \frac{r}{s}$ as the bit string 0^r1^s. (That is, write α in unary.) Then f_A^- becomes a function from bit strings (of a specified form) to $\{0,1\}$.

`Complexity Condition:` When considered as above the function f_A^- lies in the polynomial time hierarchy PH.

To be self contained we recall the definition of the polynomial time hierarchy. The functions from strings to $\{0,1\}$ are called predicates and 0 is identified with the logical value "false", while 1 is identified with "true" in this context. The complexity class P is the set of predicates computable by a Turing machine in polynomial time. The complexity class Σ_k^P (k a positive integer) consists of the predicates A given by $A(x) = \exists x_1 \forall x_2 \ldots Q x_k B(x, x_1, \ldots, x_k)$, where Q is the existential quantifier $\exists$ if k is odd, the universal quantifier $\forall$ otherwise, and B is a polynomial time predicate ($B \in P$) satisfying that in case $B(x, x_1, \ldots, x_k)$ holds then $|x_i| < |x|^c$ for $i = 1, \ldots, k$ with a constant c depending only on B. The complexity class Π_k^P is defined analogously with the role of the existential and universal quantifiers reversed. The polynomial time hierarchy PH is the union of the complexity classes Σ_k^P for $k \geq 1$.

The containments $P \subseteq \Sigma_1^P \cap \Pi_1^P$ and $\Sigma_k^P \cup \Pi_k^P \subseteq \Sigma_{k+1}^P \cap \Pi_{k+1}^P$ are self evident.

The class Σ_1^P is better known as NP. Note that the famous question whether $P = NP$ is not settled yet. If the answer is yes then all the above classes collapse to P and we have $PH = P$. We remark that in the Complexity Condition we deal with unary languages only, and the assumption that all such functions in PH are contained in P is weaker than the $P = NP$ hypothesis.

The Main Theorem of our paper is the following:

Theorem 1 *For every first order sentence A on graphs the function f_A satisfies the Very Sparse Condition, the Very Dense Condition, the Discontinuity Condition and the Complexity Condition as described above.*

Note that the validity of the first two conditions were established in [1], thus the contribution of this paper is establishing the last two conditions as stated in Theorem 16.

As mentioned at the onset, it is our belief that the converse of the Main Theorem is true. That is, if f satisfies the four conditions above then there is a first order sentence A with $f = f_A$. This we hope to return to in a sequel.

2 Rational approximations

In this section we prove elementary facts about the rational approximations $\tau(\alpha, \beta)$ and $\tau_i(\alpha)$ and the length of rationals as defined in the preceding section. While a few of these facts are used in the proof of the Main Theorem most of these observations are redundant: our motivation for this section is to better understand the statement of the Discontinuity Condition and to study our notion of length that we consider interesting on its own right.

Recall that for $\alpha \geq \beta \geq 0$ we defined $\tau(\alpha, \beta) = \sup\{k/l \leq \alpha \mid (k-1)/l \leq \beta\}$, where $k \geq 0$ and $l > 0$ must be integers.

Lemma 2 $\tau(\alpha, \alpha) = \alpha$. *For $\alpha > \beta \geq 0$ the value $\tau(\alpha, \beta)$ is a rational in the interval $(\beta, \alpha]$.*

Proof: The first statement is trivial. For the second statement choose $l_0 > 1/(\alpha - \beta)$, and set $k_0 > 0$ to the unique value with $k_0/l_0 > \beta$ and $(k_0 - 1)/l_0 \leq \beta$. The value $v_0 = k_0/l_0 \leq \alpha$ is in the set defining $\tau(\alpha, \beta)$. For any value k/l in the set $k/l \leq \beta + 1/l$ and thus it is enough to consider rationals k/l with $l \leq 1/(v_0 - \beta)$ when computing the supremum $\tau(\alpha, \beta)$. Hence it is the maximum of a finite set of rationals. □

Recall the recursive definition for $\tau_i(\alpha)$ for $\alpha \geq 0$. We set $\tau_0(\alpha) = 0$ and $\tau_{i+1}(\alpha) = \tau(\alpha, \tau_i(\alpha))$. We defined the length of a rational $\alpha \geq 0$ to be the first index i with $\tau_i(\alpha) = \alpha$.

Lemma 3 *The value $\tau_i(\alpha)$ is monotone in both i and α. For α irrational the sequence $\tau_i(\alpha)$ is a strictly increasing and tends to α. The same sequence stabilizes if α is rational, thus the length of a rational is finite.*

Proof: By the definition $\tau(\alpha, \beta)$ is monotone in both α and β and we have $\tau(\alpha, \beta) \geq \beta$ from Lemma 2. The monotonicity of $\tau_i(\alpha)$ follows.

For the rational $\alpha = k/l$ it is easy to see $\tau_i(\alpha) \geq i/l$ for $0 \leq i \leq k$. Hence the sequence $\tau_i(\alpha)$ stabilizes and the length of k/l is at most k.

For irrational α Lemma 2 implies that $\tau_i(\alpha)$ is a strictly increasing sequence of rationals. By the monotonicity for every rational $0 \leq \beta < \alpha$ we have $\tau_l(\alpha) \geq \tau_l(\beta) = \beta$ for the length $l = l(\beta)$. Thus the sequence $\tau_i(\alpha)$ must tend to α. $\square$

For $\alpha > \beta \geq 0$ let us consider $\nu(\alpha, \beta) = \sup\{k/l \mid k/l < \alpha, (k-1)/l \leq \beta\}$ where $k \geq 0$ and $l > 0$ must be integers. For $\alpha > 0$ we further define recursively $\nu_0(\alpha) = 0$ and $\nu_{i+1}(\alpha) = \nu(\alpha, \nu_i(\alpha))$.

Lemma 4 *For $\alpha > \beta \geq 0$ we have*

1. *The value $\nu(\alpha, \beta)$ is a rational in the interval (β, α).*
2. *The sequence $\nu_i(\alpha)$ is strictly increasing and tends to α.*
3. *We have $\nu_i(\alpha) = \tau_i(\alpha)$ unless α is a rational of length at most i.*
4. *For any $\gamma \in [\nu_i(\alpha), \alpha)$ one has $\tau_i(\gamma) = \nu_i(\alpha)$.*
5. *$\tau_i(\alpha) = \tau_i(\tau_j(\alpha))$ for $0 \leq i \leq j$.*
6. *For fixed $i > 0$ both $\tau_i(\alpha)$ and $\nu_i(\alpha)$ are computable from α in polynomial time (input and output are in unary).*

Proof: The first two statements is proved analogously to Lemmas 2 and 3.

For 3 notice that by definition $\nu(\alpha, \beta) = \tau(\alpha, \beta)$ unless the latter is equal to α.

For 4 notice that for $0 \leq \beta \leq \gamma < \alpha$ the set defining $\tau(\gamma, \beta)$ consist of the elements of the set defining $\nu(\alpha, \beta)$ that are at most γ. Thus if $\gamma \geq \nu(\alpha, \beta)$ we have $\tau(\gamma, \beta) = \nu(\alpha, \beta)$. The lemma follows from the recursive use of this observation.

5 follows from 3 and 4.

For 6 notice that in case $\alpha = a/b$, $\beta = c/d$ then the proof of Lemma 2 gives that the denominator of $\tau(\alpha, \beta)$ is bounded by bd. Similarly the denominator of $\nu(\alpha,\beta)$ is bounded by $bd+d$. One can compute $\tau(\alpha,\beta)$ or $\nu(\alpha,\beta)$ by considering all the denominators under these bounds one by one (notice that the input is in unary). Computing $\tau_i(\alpha)$ and $\nu_i(\alpha)$ is by the recursive definition. □

Let us remark here that using the theory of the continued fraction expansion it is not hard to compute the set $T_\alpha = \{\tau(\alpha,\beta) \mid 0 \le \beta < \alpha\}$ from the continued fraction expansion of $\alpha > 0$. T_α consists of the "best rational approximations of α from below", their continued fraction expansion are of the form $a_0 + \frac{1}{a_1+} \cdots \frac{1}{a_{2k-1}+}\frac{1}{b}$, where the expansion of α is $a_0 + \frac{1}{a_1+} \cdots \frac{1}{a_{2k}+} \cdots$ and $0 < b \le a_{2k}$. The set T_α also includes the integers $0 < b \le a_0$. The set $N_\alpha = \{\nu(\alpha,\beta) \mid 0 \le \beta < \alpha\}$ is equal to T_α if α is irrational, otherwise if the continued fraction expansion of α is $a_0 + \frac{1}{a_1+} \cdots \frac{1}{a_{2k}}$ then one can obtain N_α from T_α by removing α from the set and adding the rationals with continued fraction expansion $a_0 + \frac{1}{a_1+} \cdots \frac{1}{a_{2k}+}\frac{1}{b}$, where b is arbitrary positive integer.

Using the observations above $\tau(\alpha,\beta)$ and $\nu(\alpha,\beta)$ are computable in polynomial time even if the rationals α and β are given in *binary*. Similarly $\tau_i(\alpha)$ and $\nu_i(\alpha)$ can be computed in polynomial time from α in binary. Here i must be fixed or given in unary. (Not even this restriction on i is needed when computing $\tau_i(\alpha)$ as the length of α is bounded by $\lfloor\alpha\rfloor$ plus the length of the binary representation of α.) These observation strengthen the results in Lemma 4/6. Why we still use unary representation is explained in the proof of Lemma 14.

Recall from Section 1 that $LEN(k)$ stands for the set of non-negative rationals of length at most k.

Lemma 5 *The set $LEN(k)$ is closed and well ordered under $>$. The set of limit points of $LEN(k+1)$ is $LEN(k)$. The largest value of $LEN(k)$ less than α is $\nu_k(\alpha)$.*

Proof: The set $LEN(k)$ consists of the values $\alpha \ge 0$ with $\tau_k(\alpha) = \alpha$. The last statement follows from Lemma 4/4.

$LEN(k)$ is well ordered under $>$ since it is bounded by k and it has a largest element below any threshold.

Let $\alpha \in LEN(k)$. For any integer $j > 0$ we can choose a unique i with $i/j > \alpha \ge (i-1)/j$. We have $\tau_k(i/j) \ge \tau_k(\alpha) = \alpha \ge (i-1)/j$ thus $\tau_{k+1}(i/j) = i/j$. So $i/j \in LEN(k+1)$. The limit of these values i/j is α thus every point of $LEN(k)$ is a limit of points in $LEN(k+1)$.

It remains to prove that all limit points of $LEN(k+1)$ are in $LEN(k)$ and thus $LEN(k+1)$ is closed. We do it by induction. $LEN(0) = \{0\}$ has no limit points. Suppose the distinct points $\alpha_i \in LEN(k+1)$ tend to α. We have $\tau_k(\alpha_i) \leq \alpha_i = \tau_{k+1}(\alpha_i) \leq \tau_k(\alpha_i) + 1/j_i$ where j_i is the denominator of α_i in reduced terms. As j_i tends to infinity we have that $\tau_k(\alpha_i)$ tend to α. As $\tau_k(\alpha_i) \in LEN(k)$ by Lemma 4/5 and $LEN(k)$ is closed by induction $\alpha \in LEN(k)$. □

Examples The only number of length zero is 0. The length one numbers are of the form $1/k$ with $k > 0$ integer. The structure of $LEN(2)$ is already nontrivial. It contains, for example all rationals a/b with $b \geq a^2$. Or let us look near $\frac{1}{3}$. Clearly $\frac{1}{3} \in LEN(2)$ as it has length one, its approximation sequence being $0, \frac{1}{3}$. For $r \geq 3$ (to avoid trivialities) the values $\frac{r+1}{3r}$, $\frac{2r+1}{6r+1}$ all have length two, their approximation sequences being $0, \frac{1}{3}$, followed by the number itself. These are the only elements of $LEN(2)$ near $\frac{1}{3}$. The largest $\alpha \in LEN(2)$ with $\alpha < \frac{1}{3}$ is $\frac{5}{16}$ with approximation sequence $0, \frac{1}{4}, \frac{5}{16}$.

As in Section 1 for $\alpha \in LEN(2)$ we write α^- for $\nu_2(\alpha)$, the maximal element of $LEN(2)$ strictly smaller than α. As examples, we have $(101/300)^- = 34/101$, $(34/101)^- = 103/306$, $(103/306)^- = 69/205$, and $(1/3)^- = 5/16$.

In the rest of this section we compare our notion of length or rational numbers with related notions. None of these statements are used in the proof of the main Theorem. Our purpose here is to study the new concept of length inherent in the behaviour of first order statements on random graphs.

Let $l(\alpha)$ stand for the length of the rational $\alpha \geq 0$.

Lemma 6 *Let $\alpha > 0$ and $\tau_i = \tau_i(\alpha) = a_i/b_i$ in smallest terms for $i \geq 0$. For $i < l(\alpha)$ we have $b_{i+1} < b_i + 1/(\alpha - \tau_i)$.*

Proof: First we claim that the sequence $x_i = b_i\alpha - a_i$ is strictly decreasing until it reaches 0 at $i = l(\alpha)$. Indeed, if $b_{i+1} \leq b_i$ then $x_{i+1} = b_{i+1}(\alpha - \tau_{i+1}) \leq b_i(\alpha - \tau_i) = x_i$. Equality holds only if $\tau_{i+1} = \tau_i$. In the opposite case $b_{i+1} > b_i$ we have $\tau_{i+1} > \tau_i$ and we consider $t = (a_{i+1} - a_i)/(b_{i+1} - b_i)$. We have $t > \tau_{i+1}$ thus t cannot be in the set of which τ_{i+1} is defined to be the maximum. As $(a_{i+1} - a_i - 1)/(b_{i+1} - b_i) \leq \tau_i$ this must be because $t > \alpha$. We thus have $x_{i+1} - x_i = (b_{i+1} - b_i)\alpha - (a_{i+1} - a_i) = (b_{i+1} - b_i)(\alpha - t) < 0$ as claimed.

We have $(a_{i+1} - 1)/b_{i+1} \leq \tau_i$, hence $a_{i+1} - b_{i+1}\tau_i \leq 1$. Adding $x_{i+1} < x_i$ to this inequality one gets $b_{i+1}(\alpha - \tau_i) < x_i + 1$ and the statement of the lemma follows. □

One may find it more natural to define an approximation sequence in terms of best (one-sided) approximations with bounded denominator. We will present alternative (almost) equivalent such definitions for length.

Let us define $\alpha_i = a_i/b_i$ (in smallest terms) for a real $\alpha > 0$ with $\alpha_0 = \lfloor \alpha \rfloor$ and α_{i+1} being the maximal rational not greater than α subject to a bound on b_{i+1}. For a rational $\alpha \geq 0$ the length of this sequence is the first index i with $\alpha = \alpha_i$.

Lemma 6 motivates the $b_{i+1} < b_i + 1/(\alpha - \alpha_i)$ bound. Let $l'(\alpha)$ the length of this sequence. One may find the $b_{i+1} \leq 2/(\alpha - \alpha_i)$ condition more natural, let $l''(\alpha)$ be the length of this sequence. Finally consider the relaxed bound $b_{i+1} \leq 2/(\alpha - \alpha_i)^2$ and let $l'''(\alpha)$ be the length of this last approximation sequence.

Our next lemma states that all the above notions of length are close to each other, the same bounded sets of rationals have bounded length using any of the four variants. This shows the robustness of this notion. In particular the Discontinuity Condition can be equivalently phrased using any one of our variants. We remark however, that none of the newly definied approximation sequences have all the nice properties of the $\tau_i(\alpha)$ sequence, e.g., that $\tau_i(\alpha)$ is monotone in α.

Lemma 7 *For $\alpha \geq 0$ rational we have*

1. $l(\alpha) \geq l'(\alpha) \geq l''(\alpha) \geq l'''(\alpha)$
2. $l(\alpha) \leq \lfloor \alpha \rfloor + 2l''(\alpha)$
3. $l''(\alpha) \leq 2 \cdot 5^{l'''(\alpha)}$

Proof: The first inequality of 1 is clear from Lemma 6, the second and third inequality follows from the fact $\alpha - \alpha_i < 1$.

For 2 notice that $\tau_{\lfloor \alpha \rfloor}(\alpha) = \lfloor \alpha \rfloor$ is the starting value of the approximation sequence defining $l''(\alpha)$ and thereafter two steps of the $\tau_i(\alpha)$ sequence gets at least as high as one step of the sequence defining $l''(\alpha)$. To see this latter claim let $\beta < \alpha$ and suppose $a/b \leq \alpha$ satisfies $0 < b < 2/(\alpha - \beta)$. As $(a-2)/b \leq \beta$ and $(a-1)/b < \alpha$ we have $(a-1)/b \leq \tau(\alpha, \beta)$. From this and $a/b \leq \alpha$ follows that $a/b \leq \tau(\alpha, \tau(\alpha, \beta))$ as claimed.

Now we turn to 3. Let us fix α, let $\alpha_i = a_i/b_i$ be the approximation sequence defining $l''(\alpha)$ and let $\epsilon_i = \alpha - \alpha_i$. Recall that for $i > 0$ we defined a_i/b_i to be the largest rational not greater than α with denominator at most $2/\epsilon_{i-1}$. Consider analogously the smallest rational c_i/d_i *greater* *than* α with $d_i \le 2/\epsilon_{i-1}$. Note that b_i is strictly increasing and we have $b_i \le 2/(\alpha - \alpha_{i-1}) < b_{i+1}$. The sequence d_i is also increasing, but not necessarily strictly.

First we claim that $b_i < (4d_i)^i$. The proof is by induction on i, the base case is trivial. Suppose the claim is true for i and consider c_{i+1}/d_{i+1}. The largest acceptable value for d_{i+1} is $\lfloor 2/\epsilon_i \rfloor$, hence $c_{i+1}/d_{i+1} - \alpha \le 1/\lfloor 2/\epsilon_i \rfloor < \epsilon_i$. Thus $1/(b_i d_{i+1}) \le c_{i+1}/d_{i+1} - a_i/b_i < 2\epsilon_i)$ and $b_{i+1} \le 2/\epsilon_i < 4b_i d_{i+1} < 4d_{i+1}(4d_i)^i \le (4d_{i+1})^{i+1}$ as claimed.

As a_i/b_i and c_i/d_i are two consecutive elements in the set of rationals with denominator below a certain bound we have $c_i/d_i - a_i/b_i = 1/(b_i d_i)$. Thus $\epsilon_i < 1/(b_i d_i) < \epsilon_{i-2}/(2d_i)$ for $i \ge 2$.

Now consider the approximation sequence α'_i defining $l'''(\alpha)$. We claim that $\alpha'_i \le \alpha_{2\cdot 5^i-2}$. Here we use $\alpha_i = \alpha$ for $i \ge l''(\alpha)$. The proof is an induction on $0 \le i \le l'''(i)$. The base case is trivial. Given $\alpha'_i \le \alpha_j < \alpha$ it is enough to show that $\alpha'_{i+1} \le \alpha_{5j+7}$. We have $b_{j+2} > 2/\epsilon_j$ if $j+2 \le l''(\alpha)$. From the bound in the preceeding paragraph we get that $\epsilon_{5j+6} < \epsilon_j/((2d_{j+2})(2d_{j+4})\dots(2d_{5j+6})) \le \epsilon_j/(2d_{j+2})^{2j+3} \le 2\epsilon_j/(4d_{j+2})^{j+2} < 2\epsilon_j/b_{j+2} < \epsilon_j^2$. The value α_{5j+7} is defined to be the largest rational not exceeding α with denominator at most $2/\epsilon_{5j+6} > 2/\epsilon_j^2$ (or α if $5j+7 > l''(\alpha)$). As α'_{i+1} is defined to be the largest rational not exceeding α with denominator at most $2/(\alpha - \alpha'_i)^2 \le 2/\epsilon_j^2$ we have $\alpha'_{i+1} \le \alpha_{5j+7}$ as claimed.

Applying the last claim to $i = l'''(\alpha)$ proves statement 3 of the lemma. □

We remark that the gap between $l''(\alpha)$ and $l'''(\alpha)$ is indeed exponential. It is easy to see that $l''(\alpha - 1/n) = \Theta(\log n)$, while $l'''(\alpha - 1/n) = \Theta(\log\log n)$ for any fixed rational $\alpha > 0$.

Lemma 8 *The length $l(\alpha+\beta)$ is bounded in terms of $l(\alpha)$ and $l(\beta)$.*

Proof: We prove the statement inductively on $l(\alpha) + l(\beta)$. Let $\alpha = a/b$ and $\beta = c/d$ in smallest terms and suppose without loss of generality that $b \le d$. If $\beta = 0$ then the statement is trivial so we can consider $\beta' = \tau_{l(\beta)-1}(\beta)$. Since $\tau(\beta, \beta') = \beta$ we have $(c-1)/d \le \beta'$. By the inductive hypothesis $l = l(\alpha+\beta')$ is bounded in terms of $l(\alpha)$ and $l(\beta') = l(\beta) - 1$. We claim that $l'''(\alpha+\beta)$ is bounded by $l+1$. Indeed, the approximating sequence defining $l'''(\alpha+\beta)$ is above the $\tau_i(\alpha+\beta)$ sequence, thus if $l'''(\alpha+\beta) > l$ then the l^{th} element γ of the former sequence satisfies $\gamma \ge \tau_l(\alpha+\beta) \ge \tau_l(\alpha+\beta') = \alpha+\beta'$. Thus the next element of that sequence is $\alpha+\beta$ as its denominator bd is below the threshold $2/(\alpha+\beta-\gamma)^2 \ge 2/(\beta-\beta')^2 \ge 2d^2$. Applying Lemma 7 finishes the inductive proof. □

Let $l^*(\alpha)$ be the minimum number of reciprocals of positive integers that add up to the rational α. Notice the analogies: the same numbers have l and l^* length zero or one, the set $\{\alpha \mid l^*(\alpha) \leq k\}$ is closed and well ordered by $>$, the limit points of this set are the rationals with l^* length at most $k-1$. The analogous statements for l are stated in Lemma 5. By Lemma 8 we also know that the length $l(\alpha)$ is bounded in terms of $l^*(\alpha)$. However we believe that the converse is not true as the numbers a/b with $b > a^2$ (all in $LEN(2)$) or the numbers $1/2 + a/b$ with $b > 4a^2$ (all in $LEN(3)$) have unbounded l^* length. This would imply that the l^* analog of the Discontinuity Condition would be stronger and thus—assuming the converse of the Main Theorem—false. Notice that assuming the converse of the Main Theorem there exists a first order statement A such that f_A is not continuous at any point of $LEN(3) \cap [1/2, 1]$.

Let us finally compare our notion of length and the well known continued fraction expansion. The remark after the proof of Lemma 4 establishes a close link between the two. In particular it shows that the sum of the even indexed terms of the continued fraction expansion of a rational α is an upper bound on $l(\alpha)$. Unfortunately this does not characterize the sets of bounded length rationals. To see this, notice that for any sequence $a_2, \ldots, a_k$ and large enough a_1 the rational of continued fraction expansion $\frac{1}{a_1+}\frac{1}{a_2+}\cdots\frac{1}{a_k}$ is in $LEN(2)$ since it is of the form a/b with $b > a^2$. Thus rationals of length two can have arbitrarily long continued fraction expansion. On the other hand the rationals $1 - 1/k$ have short continued fraction expansions and unbounded lengths. On the positive side we remark that if a bounded set of rationals have bounded length continued fraction expansions and each even indexed term of those expansions are bounded by a *polynomial* of the preceding terms then the set has bounded length.

3 The α-closure

We start with definitions. Most of our notations is borrowed from or inspired by [1]. Let us recall that we consider simple undirected graphs only. The *size* of a finite graph G is (v, e) where v is the number of vertices and e is the number of edges in G. We call a pair (H, G) of a (possibly infinite) graph G and its finite subgraph H a *graph-extension.* We allow here H to have no vertices (and thus no edges) as a special case. We call H the *base.* We call the extension *trivial* if $G = H$, otherwise it is *non-trivial.* We call H' an *intermediate graph* of this extension if it is a finite subgraph of G

containing H. We call the extension finite in case G is finite. The *size* of a finite extension (H, G) is the difference between the sizes of G and H, i.e. it is (v, e) where v is the number of vertices of G not in H and e is the number of edges of G not in H.

Let $\alpha \geq 0$ be fixed. We call a non-trivial finite extension (H, G) *sparse* if its size (v, e) satisfies $v > \alpha e$. We call the same extension *dense* if it is not sparse. For technical reasons we consider the trivial extensions both sparse and dense. We call a finite extension (H, G) *rigid* if the extension (H', G) is dense for any intermediate graph H' of (H, G). We call an extension (H, G) *safe* if the extension (H, H') is sparse for any intermediate graph H' of (H, G). If α is not clear from the context we call these extensions *α-sparse*, *α-dense*, *α-rigid*, or *α-safe*.

When considering subgraphs of a common underlying graph we interpret union (respectively intersection) of graphs the standard way: taking the union (intersection) of the vertex sets and the edge sets. We can analogously define the difference $H \setminus H'$ but we must use it with more care as this operation does not yield a graph in most cases. The union of all intermediate graphs H' of an extension (H, G) such that (H, H') is α-rigid is called the *α-closure* $c_\alpha^G(H)$ of H. We suppress G in this notation if it is clear from the context. We call the subgraph H of G *α-closed* in G if $c_\alpha^G(H) = H$.

The following lemma contains simple observations.

Lemma 9 *Let H, H', H'' be finite subgraphs of a graph G. Then the following holds.*

a. If (H, H') is rigid then so is $(H \cup H'', H' \cup H'')$.

b. If (H, H') and (H', H'') are rigid then so is (H, H'').

c. If (H, H') and (H, H'') are rigid then so is $(H, H' \cup H'')$.

d. If $c_\alpha^G(H)$ is finite then it is α-closed in G and $(H, c_\alpha(H))$ is rigid.

e. If H is α-closed in G then (H, G) is safe.

Proof: a. Let H_1 be an intermediate graph in $(H \cup H'', H' \cup H'')$ and let $H_2 = H_1 \cap H'$. Notice that the sizes of $(H_1, H' \cup H'')$ and (H_2, H') are the same. The latter extension is dense as H_2 is an intermediate graph in the rigid extension (H, H'). Thus the former extension is also dense as needed.

b. Let H_1 be an intermediate graph in (H, H''). Since the graph $H_2 = H_1 \cup H'$ is an intermediate graph in the rigid extension (H', H''), the extension (H_2, H'') is dense. Using item a and that (H, H') is rigid we get that so is (H_1, H_2), so in particular it is dense. The size of (H_1, H'') is the sum of the sizes of (H_1, H_2) and (H_2, H'') thus it is also dense as needed.

c. The rigidity of $(H'', H' \cup H'')$ follows from item a, and then the claim follows from item b.

d. The second statement follows from item c. The first statement follows from the second and item b.

e. Suppose that the extension (H, G) is not safe and let H' be a minimal intermediate graph with (H, H') not sparse. Notice that $H' \neq H$. For any intermediate graph H'' of (H, H') either $H' = H''$ or (H, H'') is sparse from the minimality of H'. In either case (H'', H') is dense thus the non-trivial extension (H, H') is rigid. Hence H is not α-closed in G, giving the contrapositive of our statement. □

Suppose $\alpha > 0$ is irrational. The following statements were already identified in [3] as axioms for the almost sure theory of $G(n, n^{-\alpha})$. Here we only use that these first order statements hold almost surely in $G(n, n^{-\alpha})$.

A_H (sparsity axiom, H is a finite graph of size (v, e), $v/e < \alpha$) G does not contain a subgraph isomorphic to H.

$B^k_{H,H'}$ (safe extension axiom, (H, H') is a finite safe extension, $k > 0$ is an integer) Every isomorphism from H to a subgraph H_0 of G can be extended to an isomorphism of H' to a subgraph H_1 of G such that if for the subgraph H_2 of G the extension (H_1, H_2) is rigid of size (v, e) with $v \leq k$ then no edge of $(H_2 \backslash H_1)$ is incident to a vertex of $(H_1 \backslash H_0)$.

The sparsity axioms claim that there is no "dense" subgraph in G. The safe extension axioms claim that any finite base in G has every possible safe extension in G, moreover these extensions can be chosen not to have small rigid extensions except those of the base. In the scenario of the safe extension axiom we clearly have that $(H_2 \setminus H_1) \cup H_0$ forms a rigid extension of the base H_0. Recall that we allow that the base H in the safe extension (H, H') contains no vertices. This special case of the safe extension axiom claims that a copy of any "sparse" graph appears in G that does not have a small rigid extension.

Lemma 10 *[3] The random graph $G = G(n, n^{-\alpha})$ almost surely satisfies the sparsity and safe extension axioms.*

When using first order logic to derive consequences of the above axioms one gets the first order almost sure theory of the random graphs with edge probability $n^{-\alpha}$. Most of our lemmas are not first order though. One has to be careful when applying these axioms because no finite graph satisfies all of them. To make our reasoning simpler we deal with infinite graphs instead.

We call a graph G an α-graph if it simultaneously satisfies all the axioms above. It is clear from compactness and the Gödel completeness theorem that α-graphs exist. Due to the safe extension axion for each $k > 0$ each α-graph contains an empty subgraph on k vertices, so each α-graph is infinite. Note that by Lemma 10 any first order statement that holds for α-graphs holds almost always for the random graph $G(n, n^{-\alpha})$.

The structure of α-graphs has been studied in [2] and it was shown, for example, that for $\alpha > 1$ there is a unique countable α-graph up to isomorphism but for $\alpha < 1$ there are continuum non-isomorphic countable α-graphs. This difference is not relevant for the purposes of this paper.

While it is possible that the α-closure of a finite subgraph of an α-graph is infinite, the following lemma claims the contrary for β-closures if $\beta < \alpha$.

Lemma 11 *The β-closure $c_\beta(H)$ of a finite subgraph H of size (v, e) of an α-graph G is finite if $\beta < \alpha$. It contains at most $v/(1 - \beta/\alpha)$ vertices.*

Proof: Let H' be an intermediate graph of size (v', e') in (H, G) such that (H, H') is β-rigid. The size of this extension is $(v' - v, e' - e)$ thus $v' - v \leq \beta(e' - e)$. But the sparsity axiom gives $v'/e' \geq \alpha$. The two inequalities together imply $v' \leq v/(1 - \beta/\alpha)$. Using Lemma 9/c the statement of our lemma follows. □

We remark that the statements of Lemma 11 is not first order. For $0 < \beta < \alpha < 1$ the β-closure of the empty set in the random graph $G(n, n^{-\alpha})$ is almost always the entire graph, so any statement bounding the size of this closure fails almost always.

We need the following equivalent form of the safe extension axiom for α-graphs.

Lemma 12 *Let $\alpha > \beta > 0$ and let G be an α-graph. Let (H, H') be an α-safe extension and let f be an isomorphism from H to a subgraph H_0 of G. Then f can be extended to an isomorphism from H' to a subgraph H_1 of G. If H_0 is β-closed in G then H_1 can also be chosen β-closed.*

Proof: We use the safe-extension axiom for (H, H') and $k = \lfloor v/(1 - \beta/\alpha) \rfloor$ where v is the number of vertices in H'. We get an extension of f mapping H' to a subgraph H_1 of G as claimed. By Lemma 11 any β-rigid extension H_2 of H_1 in G must contain at most k vertices. As (H_2, H_1) is also α-rigid no edge of $(H_2 \setminus H_1)$ is incident to vertex of $(H_1 \setminus H_0)$. This implies that the sizes of the extensions (H_1, H_2) and $(H_0, (H_2 \setminus H_1) \cup H_0)$ are equal, thus the latter extension is also β-dense. If H_0 is β-closed the extensions must be trivial (Lemma 9/e), hence H_1 is β-closed. □

4 The proof

Recall the approximation $\tau(\alpha, \beta)$ for $0 \leq \beta \leq \alpha$ reals as defined by Equation 2. Recall that by Lemma 2 for $\beta < \alpha$ the value $\tau(\alpha, \beta)$ is a rational in the interval $(\beta, \alpha]$.

For notational convenience, if a graph has no edges we identify it with its set of vertices.

Lemma 13 *Let $\beta \geq 0$ be a rational, let $\alpha > \beta$, and let $\gamma = \tau(\alpha, \beta)$. Let G be an α-graph, let (H, H') be a finite extension, let f be an isomorphism from H to a subgraph H_0 of G, and let x be a designated vertex in H'. Whether there exists an isomorphism f' from H' to $c_\beta^G(H_0 \cup \{f'(x)\})$ extending f is determined by β, γ, (H, H'), x, $c_\gamma^G(H_0)$, and f. If $\gamma < \alpha$ then whether such an isomorphism exists can be decided in the second level Σ_2^P of the polynomial time hierarchy from the above inputs.*

Proof: We prove the lemma by identifying necessary and sufficient conditions for the existence of the map f'. We claim that f' exists if and only if the following two conditions are met:

1. The extension $(H \cup \{x\}, H')$ is β-rigid and
2. $c_\gamma^{H'}(H)$ has an isomorphism to a β-closed subgraph of $c_\gamma(H_0)$ extending f.

Let us start with the necessity of the conditions. Condition 1 is clearly necessary by Lemma 9/d as $(H \cup \{x\}, H')$ is isomorphic to $(H_0 \cup \{f'(x)\}, c_\beta(H_0 \cup \{f'(x)\})$ by f'. Notice that the restriction of f' to $c_\gamma(H)$ satisfies condition 2. Indeed, as $(H, c_\gamma(H))$ is γ-rigid, so is (H_1, H_0) for $H_1 = f'(c_\gamma(H))$, thus $H_1 \subseteq c_\gamma(H_0)$. If for an intermediate graph H_2 of (H_1, G) the extension (H_1, H_2) is β-rigid, then, as H_1 is contained in the image $c_\beta(H_0 \cup \{f'(x)\})$ of the isomorphism f' so is H_2. In this case the inverse image $H'' = f'^{-1}(H_2)$ in H' satisfies that $(c_\gamma(H), H'')$ is β-rigid hence $H'' = c_\gamma(H)$, since $c_\gamma(H)$ is γ-closed and $\gamma > \beta$. Thus $H_2 = H_1$ proving that H_1 is β-closed.

For the sufficiency of the two conditions we consider the extension $(c_\gamma(H), H')$ of size (v, e). By Lemma 9/d and 9/e it is γ-safe. We claim the stronger property, that it is α-safe. First we prove that it is α-sparse. We may assume $e > 0$, so we have $v/e > \gamma$ since the extension is γ-sparse. Notice that the extension $(c_\gamma(H) \cup \{x\}, H')$ is β-dense and its size is either $(v-1, e)$ or (v, e), hence $(v-1)/e \leq \beta$. These inequalities together with the definition of $\gamma = \tau(\alpha, \beta)$ implies $v/e > \alpha$ as claimed. Notice also that $1/e = v/e - (v-1)/e > \alpha - \beta$.

To prove that $(c_\gamma(H), H')$ is α-safe consider an intermediate graph H'' and let (v', e') be the size of $(c_\gamma(H), H'')$. We need to prove that this latter extension is α-sparse, so we may assume that $e' > 0$. From the γ-safe property of $(c_\gamma(H), H')$ we have $v'/e' > \gamma$. In case $(v'-1)/e' \leq \beta$ we can again use the definition of $\gamma = \tau(\alpha, \beta)$ to derive that $v'/e' > \alpha$. In the opposite case we have $v'/e' = (v'-1)/e' + 1/e' > \beta + 1/e > \alpha$. Thus $(c_\gamma(H), H'')$ is α-sparse in both cases as claimed.

Let f'' be the isomorphism claimed in condition 2 and let H_1 be its image in $c_\gamma(H_0)$. By assumption H_1 is β-closed in $c_\gamma(H_0)$, so by Lemma 9/d it is β-closed in G. We apply Lemma 12 for the α-safe extension $(c_\gamma(H), H')$ and the isomorphism f''. By the lemma f'' can be extended to an isomorphism f' of H' to a β-closed subgraph H_2 of G. Here $(H_0 \cup \{f'(x)\}, H_2)$ is isomorphic to $(H \cup \{x\}, H')$, hence β-rigid by condition 1. Thus $H_2 \subseteq c_\beta(H_0 \cup \{f'(x)\}$. Since H_2 is β-closed we have equality in the last formula proving the sufficiency of our two conditions.

Note that the two conditions are formulated in terms of our inputs only. To prove the claim of the lemma about computability notice that by Lemma 11 $c_\gamma(H_0)$ is finite. Condition 1 can be checked in Π_1^P as we only have to check that (H'', H') is β-dense for all intermediate graphs H'' of $(H \cup \{x\}, H')$. To check condition 2 we have to guess $c_\gamma(H)$ and the isomorphism to a subgraph of $c_\gamma(H_0)$ and check if its image is β-closed. We also have to verify that $c_\gamma(H)$ is guessed correctly by checking if it is γ-closed and extends H in a γ-rigid way. All this can be done in Σ_2^P. $\square$

Although immaterial in the present proof it would be nice to strengthen the above lemma by replacing Σ_2^P by NP. Checking the second condition is closely related to the subgraph isomorphism problem indicating perhaps that the our problem is also NP-hard.

Let us recall that the variables of first order statements on a graph range over the vertices of the graph. First order formulae are built from the atomic formulae of the form $x = y$ or $x \sim y$ with the use of logical connectives $\wedge$, $\vee$ and $\neg$ and quantifiers $\exists x$ and $\forall x$. The (open) formulae without quantifiers are called *propositional formulae.* A *prenex formula* is a propositional formula prefixed with a sequence of quantifiers for distinct variables. The variables in the quantifiers are the closed variables, the remaining variables in the formula are the open variables. A formula with no open variables is called a closed formula or a statement. Recall that every statement is equivalent with a prenex statement.

Recall the approximating sequence $\tau_i(\alpha)$ as defined in the first section by $\tau_0(\alpha) = 0$ and by the recursion given in Equation 3. By the *unary* representation of the rational $r \geq 0$ we mean a bit string $0^k 1^l$ with $k/l = r$.

Lemma 14 *Let A be a fixed prenex formula with i closed and j open variables. Suppose $\alpha > 0$ satisfies $\tau_i(\alpha) < \alpha$. Whether A holds in the α-graph G for the vertices $X_1, \ldots, X_j$ is determined by A, the graph $c_{\tau_i(\alpha)}(\{X_1, \ldots, X_j\})$ and its vertices $X_1, \ldots, X_j$ and by the number $\tau_i(\alpha)$. Deciding from these inputs if a fixed formula A holds lies within the $(i+1)^{st}$ level Σ^P_{i+1} or Π^P_{i+1} of the polynomial time hierarchy if the rational $\tau_i(\alpha)$ is given in unary.*

Proof: The proof is by induction on i. For $i = 0$ we have a propositional statement, this is clearly determined by the subgraph of G spanned by $X_1, \ldots, X_j$ which is exactly $c_0(\{X_1, \ldots, X_j\})$. The decision is in constant time.

If $i > 0$ A has the form $\exists x B$ or $\forall x B$. By the De Morgan law we can write the negation of the formula of the second type as a formula of the first type, thus it is enough to consider the case $A = \exists x B$. Here B is a prenex formula of $i - 1$ closed and $j + 1$ open variables. (We suppose x appears in B.) By the inductive hypothesis, whether B holds in G for the vertices $X_1, \ldots, X_j$ and an extra vertex X for the variable x is determined by the $\tau_{i-1}(\alpha)$ closure of these vertices and the approximation $\tau_{i-1}(\alpha)$. Notice here that $\tau_{i-1}(\alpha) = \tau_{i-1}(\tau_i(\alpha))$, thus the previous approximation of α is computable from the next in polynomial time (Lemma 4).

We can thus determine if A holds in G for $X_1, \ldots, X_j$ by guessing the graph $c_{\tau_{i-1}(\alpha)}(\{X_1, \ldots, X_j, X\})$ up to an isomorphism fixing $X_1, \ldots, X_j$, and X, and checking if i) with this closure B holds and ii) there exists a vertex X in G for which the closure is guessed correctly. Item i can be checked within the i^{th} level of the polynomial time hierarchy by induction, while item ii can be checked within Σ^P_2 by Lemma 13. To finish the proof it remains to bound the size of the $\tau_{i-1}(\alpha)$ closure to be guessed in the beginning. By Lemma 11 it has at most $(j+1)/(\alpha - \tau_{i-1}(\alpha)) < (j+1)/(\tau_i(\alpha) - \tau_{i-1}(\alpha))$

vertices. We assumed that the approximations of α are given in unary to ensure that this is polynomial in the input size. $\Box$

Theorem 15 *For any first order statement A and any irrational $\alpha > 0$ either A holds for all α-graphs or it holds from none. There exists an integer $i > 0$ for A such that whether A holds for an α-graph can be decided within the polynomial time hierarchy PH from $\tau_i(\alpha)$ given in unary.*

Proof: First write A in prenex form of some i closed variables, then apply Lemma 14. By the sparsity axiom the $\tau_i(\alpha)$ closure in an α-graph of the empty set is empty since $\tau_i(\alpha) < \alpha$. Thus by Lemma 14 whether A holds is independent of the graph since it can be computed in the polynomial time hierarchy PH solely based on $\tau_i(\alpha)$. $\Box$

We remark here that the integer i found by the proof is the number of variables in a prenex formula equivalent to A. We used prenex formulae for convenience only, a more careful analysis shows that i can be chosen to be the *quantifier depth* of A i.e., the length of the longest nested sequence of quantifiers in A.

Theorem 15 establishes the 0-1 law for the first order sentences on the random graphs $G(n, n^{-\alpha})$ for any irrational $\alpha > 0$. This has already been proven in [1]. Our result gives more in telling how the validity of the statement depends on α as claimed in the next theorem.

Recall that for a first order statement A on graphs we defined the function f_A on the positive irrationals by Equation 1. We identified four conditions on this function in the Section 1, among them the Discontinuity Condition and the Complexity Condition.

Theorem 16 *For a first order statement A on graphs the function f_A is a 0-1 valued function satisfying the Discontinuity Condition and the Complexity Condition.*

Proof: By Theorem 15 f_A is 0-1 valued and one can find an integer $i > 0$ such that $f_A(\alpha) = f_A(\alpha')$ if $\tau_i(\alpha) = \tau_i(\alpha')$ (here α and α' are positive irrationals). As the value of $\tau_i(\alpha')$ is constant in a small interval around any positive α unless $\tau_i(\alpha) = \alpha$ (Lemma 4) the Discontinuity Condition follows.

Recall the definition f^- as given in Equation 4. By Lemma 4/4 we have $\tau_i(\alpha') = \nu_i(\alpha)$ for $\alpha > 0$ and $\alpha' \in [\nu_i(\alpha), \alpha)$. Thus by Lemma 15 the limit $f_A^-(\alpha)$ exists and it is computable in PH from $\nu_i(\alpha)$ given in unary. Thus by Lemma 4/6 $f_A^-(\alpha)$ can be computed in the polynomial time hierarchy from the rational $\alpha > 0$ given in unary. $\Box$

This last Theorem establishes the Main Theorem as the Very Sparse Condition and the Very Dense Condition have already been established in [1].

We remark here that even assuming that the four conditions in the Main Theorem characterize the functions f_A a subtle issue with respect to the spectrum of first order sentences remain unsolved. The spectrum Spec(A) of a first order statement A is defined in [1] as the set of values $a > 0$ for which there is no value $\epsilon > 0$ and $\delta \in \{0,1\}$ such that $\Pr[G(n,p) \models A]$ tends to δ as n goes to infinity with $n^{-a-\epsilon} < p < n^{-a+\epsilon}$. While all the points of discontinuity of the function f_A are among the spectrum Spec(A) the converse is false. Consider the first order statement that G has a unique 4-clique. The function f_A is constant zero as $G(n,n^{-\alpha})$ has no 4-cliques for $\alpha > 2/3$ and it has many 4-cliques for $\alpha < 2/3$. The probability $\Pr[G(n,n^{-2/3}) \models A]$ tends to a positive limit less than one, thus $2/3 \in$ Spec(A) but this is not noticeable from the function f_A. The techniques of this paper give the following stronger version of the Discontinuity Condition:

The spectrum Spec(A) of a first order statement A of quantifier depth i consists of rationals of length at most i.

We do not know what modification of the Complexity Condition one needs to characterize the spectrum of first order sentences.

References

[1] S. Shelah, J. Spencer, Zero-One Laws for Sparse Random Graphs, *J. Amer. Math. Soc.* **1** (1988) 97-115.

[2] J. Spencer, Countable Sparse Random Graphs, *Random Structures and Algorithms* **1** (1990) 205-214.

[3] J. Spencer, Sparse Random Graphs: A Continuum of Complete Theories *in* Sets, Graphs and Numbers (eds.: G. Halász, L. Lovász, D. Miklós, T. Szőnyi), Colloquia Math Soc. János Bolyai, 60, Budapest 1991, 679-690.