On Zero-One and Convergence Laws for Graphs Embeddable on a Fixed Surface

Albert Atserias[1]

Universitat Politècnica de Catalunya, Barcelona, atserias@cs.upc.edu

Stephan Kreutzer[2]

Technical University Berlin, stephan.kreutzer@tu-berlin.de

Marc Noy[3]

Universitat Politècnica de Catalunya, Barcelona, marc.noy@upc.edu

―――― **Abstract** ――――

We show that for no surface except for the plane does monadic second-order logic (MSO) have a zero-one-law – and not even a convergence law – on the class of (connected) graphs embeddable on the surface. In addition we show that every rational in [0,1] is the limiting probability of some MSO formula. This strongly refutes a conjecture by Heinig et al. (2014) who proved a convergence law for planar graphs, and a zero-one law for connected planar graphs, and also identified the so-called gaps of [0,1]: the subintervals that are not limiting probabilities of any MSO formula. The proof relies on a combination of methods from structural graph theory, especially large face-width embeddings of graphs on surfaces, analytic combinatorics, and finite model theory, and several parts of the proof may be of independent interest. In particular, we identify precisely the properties that make the zero-one law work on planar graphs but fail for every other surface.

1 Introduction

We consider classes of labelled graphs with the uniform distribution on graphs with a fixed number of vertices. Let S be a closed compact surface, and let $\mathcal{G}_S$ be the class of graphs embeddable on S. Heinig et al. [9] studied limiting probabilities of first order (FO) and monadic second order (MSO) properties of graphs in the classes $\mathcal{G}_S$. They showed that a convergence law in FO holds for all surfaces S, and that a convergence law holds in MSO when S is the sphere, so that $\mathcal{G}_S$ is the class of planar graphs. They also showed that a zero-one law in FO holds for *connected* graphs in $\mathcal{G}_S$, and a zero-one law in MSO for connected planar graphs. They conjectured that these results extend to MSO properties on

45th International Colloquium on Automata, Languages, and Programming (ICALP 2018).
Editors: Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella;
Article No. 116; pp. 116:1–116:14

Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

an arbitrary surface. We strongly refute these conjectures. For each surface S other than the sphere we construct an MSO formula φ such that the probability that φ is satisfied in $\mathcal{G}_S$ does not converge when the number of vertices n tends to infinity, not even on connected graphs. In addition, telling whether the probability of a given MSO formula φ converges is an undecidable problem. We also show that every rational number in [0,1] is the limiting probability of some MSO formula for connected graphs in $\mathcal{G}_S$.

We sketch next the main ingredients in the proofs of our results. We fix a surface S different from the sphere. Embeddings of graphs on surfaces can be defined in purely combinatorial terms. If S is orientable, an embedding of a graph on S is given by a rotation system consisting, for each vertex v, of a cyclic orientation of the edges incident with v. If S is non-orientable one has to consider also signed edges [10]. Given a graph G embedded in S the *face-width* of the embedding is the minimum number of intersections of G with a non-contractible curve in S. In what follows, we say that a graph property holds with high probability (w.h.p.), or asymptotically almost surely (a.a.s.), if it holds with probability tending to 1 as n tends to infinity. It is known [10] that if the face-width of a 3-connected graph G is at least $2g$, where g is the genus of G (orientable or not), then G has a unique embedding in S. We also need the fact that the face-width of a random 3-connected graph G embedded on S is $\Omega(\log n)$ w.h.p. [1].

It is shown in [2] that w.h.p. a random graph G in $\mathcal{G}_S$ has a unique non-planar 3-connected component C. Since planarity is MSO expressible and 3-connected components are MSO definable, so is C. Using the fact that w.h.p. the face-width of C is large, we show the existence of an MSO definable grid structure M in G of size $\Omega(\log \log n)$. This is obtained starting with a non-contractible cycle, which is MSO definable, and extending it to a grid structure. Inspired by the capacity of MSO to emulate Turing machine computations on grid graphs, we are then able to define an MSO formula φ expressing the property that $\log^\dagger |M|$ is in $\{0, 1, 2, 3\}$ modulo 8, where $\log^\dagger n$ is a variant of $\log^* n$ (see Section 5). Given that $\log \log n \leq |M| \leq n$, the value of $\log^\dagger |M|$ modulo 8 oscillates and, as a consequence, the probability that φ holds does not converge as n goes to infinity.

The non-converging formula φ, combined with the abovementioned capacity of MSO to emulate Turing machine computations on grid graphs, also gives the undecidability of the decision problem for converging probabilities. For each Turing machine M, we find an MSO formula φ_M that simulates partial runs of M along the definable growing grids. The formula φ_M has asymptotic probability 1 if M halts, and asymptotic probability 0 otherwise, and the probability of the conjunction $\varphi \wedge \varphi_M$ is thus converging if and only if M halts.

To prove that each rational in $[0, 1]$ is the limiting probability of some MSO formula, we use the following facts. (a) A 3-connected graph in a surface of genus g has a spanning tree with maximum degree at most $4g$ [7]. (b) For the class $\mathcal{G}_S$, every property in MSO_2 (quantification over vertices and edges) is also expressible in MSO [5]. (c) The size X_n of the unique non-planar component obeys a limit local law related to a stable law [8]. Consider now a random graph G in $\mathcal{G}_S$ with n vertices, let C be the unique non-planar component, and let $X_n = |C|$. From properties (a) and (b) it follows that, for each integers a and b, we can express in MSO the property $\varphi_{a,b}$ that X_n is equal to a modulo b. Using property (c) we show that the probability that $\varphi_{a,b}$ is satisfied tends to $1/b$. Now given any rational non-negative $a/b \leq 1$, the property that X_n is less than a modulo b tends to a/b.

The key fact that makes MSO properties of graphs on non-planar surfaces so different from the plane is that w.h.p. there is a unique *non-planar* 3-connected component, which also happens to be the only one of linear size. For random planar graphs there is a unique 3-connected component of linear size as well, but it is indistinguishable in MSO from the smaller ones.

2   Graphs and Surfaces

In this section we present the background from graph theory and surface topology we need for our main result. We refer to [10, 6] for background on graphs and surfaces. Our notation on surfaces and embeddings follows [10].

Graphs.   All graphs in this paper are finite, undirected, and simple, i.e. no parallel edges or self-loops. We denote the vertex and edge set of a graph G by $V(G)$ and $E(G)$, resp. For $d \geq 0$, a graph G is *d-degenerate* if every subgraph $H \subseteq G$ contains a vertex of degree at most d.

We need the concept of 3-connected components of a graph G. For $k > 0$, a graph G is *k-connected* if $|V(G)| > k$ and $G - S$ is connected for all $S \subseteq V(G)$ with $|S| < k$. A *connected component* of G is an inclusion-wise maximal connected subgraph $C \subseteq G$. A *2-connected component* of G is an inclusion-wise maximal 2-connected subgraph $C \subseteq G$. For defining *3-connected components* of G we need some preparation.

Let G be a 2-connected graph. A *separator of order* 2 in G is a set $X = \{u, v\}$ of distinct vertices such that $G - X$ is not connected. Let C be an inclusion-wise maximal subgraph of G such that for all 2-separations X of G there is one connected component of $G - X$ which contains all vertices of $C - X$. For any such 2-separation X of G, if $X \subseteq V(C)$ then we add an edge between the two vertices of X in C if it is not already there. This produces a new graph $\widetilde{C}$. For a 2-connected graph G, the graphs $\widetilde{C}$ obtained in this way which are not cycles are called the 3-connected components of G. Tutte proved that every 2-connected graph has a decomposition into a tree whose nodes are cycles or 3-connected components.

Surfaces and Graph Embeddings.   We now present some fundamental properties of surfaces and embedded graphs. We refer to [10] for background.

A *surface* is a compact connected Hausdorff topological space in which every point has a neighbourhood homeomorphic to the plane. A surface can be constructed from the sphere by cutting a number of holes and pairs of holes into the sphere, each homeomorphic to an open disc. Every pair of holes is then closed by adding a *handle* – an open cylinder – connecting the boundary of the holes. The remaining holes are closed by adding a *crosscap*, that is, by identifying each point on the boundary of the hole with the corresponding point on the opposite side. The surface classification theorem shows that every surface is homeomorphic to a surface constructed in this way. Any surface obtained in this way that includes a crosscap is called *non-orientable*, otherwise it is called *orientable*. Our main result holds for orientable and non-orientable surfaces. Due to space constraints, we only explain the orientable case. The surface obtained from the sphere by adding k handles is denoted by $\mathbb{S}_k$. The number k is called the *genus* of $\mathbb{S}_k$.

Let $S = \mathbb{S}_k$ be a surface. The way we constructed S implies that we can reduce it to the plane by taking for each handle a closed curve that goes around the handle and cut the surface along the curves, closing the appearing holes by disks. This way, every handle is cut open and the resulting surface is homeomorphic to the plane. We call curves which cut a handle in this way *noncontractible*. There are two types of non-contractible curves: if we cut along a curve, we may either disconnect the surface or not. Curves which do not disconnect the surface when we cut along them are called *non-separating*.

Following [10], a graph G' is *embedded* on a surface S if its vertices are distinct points on S and every edge e of G' with endpoints u, v is a simple closed arc connecting u and v in S such that the interior of e is disjoint from other edges and vertices of G'. A graph G is

embeddable on S if it is isomorphic to an embedded graph G' on S. Let Π be an embedding of a graph G in S. The connected components of $S - \Pi$ are called the *faces* of Π.

An embedding of a graph in a surface S can be uniquely represented by an object called an *embedding scheme* or *rotation system*. If v is a vertex of G then the embedding of G embeds the edges incident with v in some order in clockwise orientation around v. That is, for any v the embedding defines a cyclic permutation of the incident edges, or a linear order on the edges that have v as an endpoint. We call this order the *clockwise order* around v. A *rotation system* is a set of clockwise orderings containing one order for every vertex.

Such a rotation system uniquely determines an embedding of G in S. Given a rotation system π of a graph G, we can construct the facial cycles of the embedding as follows. Let v be a vertex and let $e = \{u, v\}$ be an incident edge. Then v and e determine a walk in G where we start at v, follow the edge e to its other endpoint and then proceed with the next edge in clockwise order until we return to v. In this way we obtain exactly the facial cycles of the embedding. In particular, we can represent any face of the embedding by orienting an edge in its boundary cycle, which fixes the start vertex v and the edge e to choose first. This will be used in the MSO definition in the next section.

One of the main challenges in proving our main result is that we need to define MSO-formulas that encode in a given graph an embedding of it in a fixed surface S. Whitney proved that a planar graph G that is 3-connected has a unique embedding into the plane. In [3], Courcelle proved that there are MSO-formulas defining the rotation system for planar graphs.

To prove that rotation systems are also definable for graphs on other surfaces we will reduce the problem to the planar case as follows. As mentioned above, any surface can be reduced to the plane by cutting the surface along a finite set of closed curves each of which cuts a handle. We can generalize this concept of *cutting along a cycle* to cycles in embedded graphs.

In the sequel, let S be a surface. If we refer to a Π-*embedded graph* G we implicitly define Π to be an embedding of the graph G on S. Let $C = (v_0 e_1 v_1 e_2 .... v_k = v_0)$ be a cycle in G. For $i > 0$, let L_i be the set of edges incident to v_i which in the clockwise ordering around v_i appear after e_i but before e_{i+1} the R_i be the set of edges appearing after e_{i+1} and e_i. The edges in any L_i are called the *left* edges of the cycle C and the edges in any R_i are called the *right* edges of C.

▶ **Definition 2.1.** Let $C = (v_0 e_1 v_1 \ldots v_k)$ be a surface nonseparating cycle of a Π-embedded graph G. Let G' be the graph obtained from G by replacing C by two isomorphic copies $C_r = (v_0^r e_1^r v_1^r \ldots v_k^r)$ and $C_l = (v_0^l e_1^l v_1^l \ldots v_k^l)$ such that all edges $e = \{u, v_i\}$ on the right of C are replaced by edges $\{u, v_i^r\}$ incident to the corresponding vertices on C_r and all edges $e = \{u, v_i\}$ on the left of C are replaced by edges $\{u, v_i^l\}$ incident to the corresponding vertices on C_l. We say that G' is obtained from G by *cutting along the cycle C*. The embedding Π defines an embedding Π' of G' in the surface obtained from S by cutting along C and closing the resulting holes by discs in the obvious way. Note that the two copies C_1, C_2 of C are now facial cycles. We now add two new vertices f_1, f_2 such that f_i has an edge to every vertex of C_i. This means that f_i is drawn in the face bounded by C_i. We call the resulting graph the *augmented graph obtained by cutting along C*.

This motivates the following definition.

▶ **Definition 2.2** (Planarizing set of cycles). Let G be a Π-embedded graph. A *planarizing set of cycles* is a set $C_1, \ldots, C_k$ of pairwise disjoint cycles in G such that cutting along all cycles $C_1, \ldots, C_k$ results in a connected graph embedded in the sphere.

We will see below that for certain graphs embedded on a surface such sets of cycles always exist. For this, we need the concept of face-width.

▶ **Definition 2.3.** Let Π be an embedding of a graph G on S. The *face-width* $\mathrm{fw}(G, \Pi)$ of G is the smallest number k such that S contains a noncontractible closed curve that intersects G in k points, or ∞ if no noncontractible curve exists (i.e. S is the plane).

We prove next that the connectivity between two Π-noncontractible and Π-nonseparating cycles is at least as high as the face-width of Π.

▶ **Lemma 2.4.** *Let G be a 2-connected Π-embedded graph and let C and C' be two disjoint Π-noncontractible, Π-nonseparating cycles. Let $2 \leq k = \mathrm{fw}(G, \Pi) < \infty$. Then there are at least k pairwise vertex disjoint paths linking C and C'. Furthermore, if G is 3-connected with $3 \leq \mathrm{fw}(G, \Pi) < \infty$ and C is a Π-nonseparating cycle then be the augmented graph $\widetilde{G}$ obtained from G by cutting along C is 3-connected.*

See [10, Theorem 5.11.2] for a proof of the following theorem.

▶ **Theorem 2.5.** *Let G be a graph that is Π-embedded in a surface with Euler genus g and let d be a positive integer. If $\mathrm{fw}(G, \Pi) \geq 8(d+1)(2^g - 1)$, then G contains a planarizing collection of induced cycles $C_1, \ldots, C_k$, for some $g/2 \leq k \leq g$, such that the distance between C_i and C_j, $i \neq j$, is at least d.*

The previous theorem and the lemma preceding it show that starting from a graph that is 3-connected and has an embedding of face-width at least 3, we can reduce it to a 3-connected planar graph by cutting along a constant number of cycles. Combining this with Whitney's theorem of unique embeddings of 3-connected planar graphs implies the following result which will be important later.

▶ **Theorem 2.6.** *Let G be a 3-connected graph that is Π-embedded in a surface with Euler genus g such that $3 \leq 8(d+1)(2^g - 1) \leq \mathrm{fw}(G, \Pi) < \infty$. Let G' be the augmented graph obtained by cutting along a set of planarizing cycles. Then the facial cycles of Π are precisely the facial cycles of the unique plane embedding of G'. Moreover, an embedding scheme which is equivalent to Π can be deduced from the rotation system of the plane embedding of G'.*

The following definition captures the abstract properties of planarizing sets of cycles.

▶ **Definition 2.7.** Let G be a graph and let $k \geq 0$. A *potential system of planarizing cycles of order k* is a sequence $\big((C_i, L_i, R_i)\big)_{i=1}^k$ such that $C_1, \ldots, C_k$ are pairwise vertex disjoint cycles in G and L_i, R_i form a partition of the set of edges $e \notin E(C_i)$ incident with a vertex on C_i, for all i.

Note that the procedure of cutting along a cycle in a graph G and the augmented graph obtained from G in this way as defined above can be applied to any cycle C with a given partition L, R of the edges $e \notin E(C)$ incident with a vertex on C. Of course it may not always lead to the intended effect, e.g. if the sets L and R of left and right edges are chosen wrongly. But in any case it will produce a graph G' and if G and Π satisfy the conditions of Theorem 2.5, then G' will be the augmented graph obtained by cutting along the cycles $C_1, \ldots, C_k$.

Let G' be the graph obtained in this way. We call G' the graph *obtained from G by cutting along $\mathcal{S}$*. If G' is planar and Π' is a plane embedding of G' then we can obtain an embedding Π'' of G on some surface as follows. Let $f_1, \ldots, f_k$ be the extra vertices in G'. If we delete $f_1, \ldots, f_k$ from G' and Π' then we obtain a plane embedding Π'' which, for each i,

has two facial cycles corresponding to the two copies of C_i. We cut two holes, each having a copy of C_i as boundary cycle, and then identify the corresponding points on the two copies of C_i in the obvious way. In this way we obtain a surface S defined by an embedding Π''' of G. We call S, Π''' the surface and embedding *obtained from G' and Π' by* gluing *along $\mathcal{S}$*.

The results and constructions proved in this subsection imply the next theorem.

▶ **Theorem 2.8.** *Let S be a surface of genus g and let G be a 3-connected graph embedded in S by an embedding Π such that $8(d+1)(2^g - 1) \leq \text{fw}(G, \Pi) < \infty$.*

1. *Then, there exists a potential system of planarizing cycles $\mathcal{S}$ of order at most g such that the graph G' obtained from G by cutting along $\mathcal{S}$ is 3-connected and planar, and if Π' is the (uniquely determined) plane embedding of G', then S and Π are the surface and the embedding obtained from G' and Π' by gluing along $\mathcal{S}$.*

2. *Furthermore, for every potential system of planarizing cycles $\mathcal{S}$ of order at most g such that the graph G' obtained from G by cutting along $\mathcal{S}$ is 3-connected and planar, if Π' is the (uniquely determined) plane embedding of G', then S and Π are the surface and the embedding obtained from G' and Π' by gluing along $\mathcal{S}$.*

This theorem is the main tool for defining embeddings in monadic second-order logic later on. In the next section we exploit face-width for finding grids of controlled size.

Grid-like Structures in High-Face Width Embeddings.   In this section we establish some graph theoretical properties of embedded graphs that will allow us to define grids in embedded graphs whose order is proportional to the face-width.

▶ **Definition 2.9.** Let G be a 2-connected Π-embedded graph such that $\text{fw}(G, \Pi) \geq 2$ and let $F = F(G, \Pi)$ be the set of Π-faces. The *vertex-face graph* is the bipartite graph $\Gamma = \Gamma(G, \Pi)$ with vertex set $V(G) \cup F(G, \Pi)$ and an edge between $u \in V(G)$ and $f \in F(G, \Pi)$ if u is contained in the facial cycle bounding f.

Note that any closed curve on a surface corresponds to a cycle in the vertex-face graph in the obvious way. From now on, we will therefore consider cycles in the $\Gamma(G, \Pi)$. The length of such a cycle is the number of faces (or vertices) on it.

▶ **Lemma 2.10** (Prop. 5.5.10 of [10])**.** *Let G be a Π-embedded graph such that $2 < \text{fw}(G, \Pi) < \infty$, and let $k := \lfloor \frac{\text{fw}(G, \Pi)}{2} \rfloor - 1$. Let v be a Π-face and let $B_0(v)$ be the Π-boundary walk of v. For $i > 0$ we define $B_i(v)$ as the union of $B_{i-1}(v)$ and all Π-facial walks that have a vertex in $B_{i-1}(v)$. Then there exist $k + 1$ disjoint Π-contractible cycles $C_0, \dots, C_k$ such that for all $i = 0, 1, \dots, k$, $C_i \subseteq \partial B_i(v)$ and $B_i(v) \subseteq Int(C_i)$.*

We are now ready to prove the graph theoretical properties we will use later to show that any 3-connected graph embedded by a face-width k embedding, for some $3 \leq k < \infty$, contains a $\frac{k}{4} \times \frac{k}{4}$-grid which, moreover, is *controlled* by a noncontractible cycle.

Let Π be an embedding of a graph G on a surface S of Euler genus g such that $\text{fw}(G, \Pi) \geq 3$. Let $\Gamma = \Gamma(G, \Pi)$. Let $k := \lfloor \frac{\text{fw}(G, \Pi)}{2} \rfloor - 1$.

▶ **Definition 2.11.** Let C be a Π-nonseparating, Π-non-contractible cycle of Γ of minimal length. Hence, the number of vertices and faces on C is exactly $\text{fw}(G, \Pi)$. Let v be a face on C. A set $\mathcal{P} = \{P_0, \dots, P_k\}$ of pairwise vertex disjoint cycles is *controlled by C and v*, if for all $0 \leq i \leq k$,

- P_i intersects C in exactly two vertices p_i, p'_i,
- for all $1 \leq i \leq k$, p_i and p_{i-1} and also p'_i and p'_{i-1} have a common neighbour on C, which is a face, and p_0 and p'_0 are the neighbours of the face v on C, and
- if $i < j$ then P_i is contained in the component of $G - P_j$, called the *interior* of P_j, that contains v.

Let C and v be as before and let $\mathcal{P} = \{P_0, \ldots, P_k\}$ be a set of pairwise vertex disjoint cycles controlled by C and v. Then either there is exactly one face $\bar{v}$ on C not in the interior of P_k or there is exactly one vertex $\bar{v}$ on C which is neither in the interior of P_k nor on P_k itself. We call the node $\bar{v}$ in the previous claim the *opposite node of v on C* and denote it by $\bar{v}$.

Let $C, v, \bar{v}$ and $\mathcal{P}$ be as before. Let a be a face on C and let $\mathcal{Q} = \{Q_1, \ldots, Q_k\}$ be a set of pairwise disjoint cycles controlled by C and a. Let $\bar{a}$ be the opposite node of a on C.

Let $s := \lceil \frac{k+1}{2} \rceil$ and let $s' := k - s$. If $\mathrm{fw}(G, \Pi)$ is odd then a is a face of facial distance s from v on C if, and only if, the cycle $Q \in \mathcal{Q}$ which contains the vertex $\bar{v}$ has v on its exterior but v is adjacent to a vertex on Q. If $\mathrm{fw}(G, \Pi)$ is even then a is a face of facial distance s from v on C if, and only if, v and $\bar{v}$ are adjacent to vertices on the same two cycles in $\mathcal{Q}$. If a and v satisfy these conditions, then we say that v and a *match*.

This observation will allow us to define $a, \bar{a}, \bar{v}$ from C and v. What is left is to give a topological condition for C to be a noncontractible cycle of minimal length. But this can easily be done using Lemma 2.4. Finally, it can be shown that $C, v, \bar{v}, a, \bar{a}$ uniquely determine a grid structure in $C \cup \mathcal{P} \cup \mathcal{Q}$ of size s'. The previous claims together establish the following theorem.

▶ **Theorem 2.12.** *Let C be a noncontractible, nonseparating cycle of length $\mathrm{fw}(G, \Pi)$ and let v be a face on C. Let $\mathcal{P}$ be a set of pairwise disjoint cycles controlled by C and v and let $\bar{v}$ be the node on the opposite of v on C. Let x be a face on C and let $\mathcal{Q}$ be a set of pairwise disjoint cycles controlled by C and x and let $\bar{x}$ be the node opposite of x on C. Finally, suppose v and x match as defined in the previous claim. Then $C, v, x, \bar{x}, \bar{v}$ determine an $s' \times s'$ grid.*

3 Monadic Second-Order Logic

In this section we introduce monadic second-order logic (MSO), and develop the MSO definability of embedding schemes and grids in large face-width embeddings.

Logic. A *vocabulary* is a set of relation symbols with associate arities. If L is a vocabulary, a *finite L-structure M* is given by a finite *domain* or *universe* $D(M)$ and a relation $R(M) \subseteq M^r$ for each relation symbol $R \in L$ of arity r

The class of formulas of monadic second-order logic (MSO) is the smallest class of formulas that contains the atomic formulas and is closed under negations, conjunctions, disjunctions, and existential and universal quantification of individual and set variables. An individual variable ranges over the domain, and a set variable ranges over the subsets of the domain. If X is a set variable and x is an individual variable, then $X(x)$ is the atomic formula that says that x is in the set X. Given a structure M, the relation defined by a formula $\varphi(\bar{x}, \bar{X})$ with individual and set free-variables $\bar{x}$ and $\bar{X}$, respectively, is the set of pairs of tuples $(\bar{a}, \bar{A})$ such that M makes $\varphi(\bar{a}, \bar{A})$ true. A formula without free variables is called a *sentence*.

Logic on Graphs, MSO₁ and MSO₂. There are two natural encodings of graphs G as structures. In the *standard encoding* the vocabulary has a single binary relation symbol E, the domain of the structure is $V(G)$, and the binary relation E is interpreted by $E(G)$. In the *incidence encoding* the vocabulary has two unary relation symbols V and E, one binary relation symbol I, the domain is $V(G) \cup E(G)$, and I is the incidence relation between vertices and their incident edges. The unary relations V and E are interpreted by $V(G)$ and $E(G)$. Whereas for FO it is irrelevant which encoding is used, the two encodings behave

differently for MSO: on the standard encoding the set quantifiers range over sets of vertices whereas on the incidence encoding they range also over sets of edges, e.g., over paths. MSO on the standard encoding is often referred to as MSO_1 and on the incidence encoding it is referred to as MSO_2. For graphs of bounded genus, and even more generally, for classes of graphs that are p-degenerate for some fixed $p > 0$, the two logics MSO_1 and MSO_2 are equivalent: see Theorem 5.22 in [5]. Since we only consider classes of graphs of bounded genus we will assume from now on that graphs are given by their incidence encoding.

MSO-definitions and interpretations. For an integer $k \geq 1$, an MSO-definition of L-structures of *order k* is a collection of formulas $\varphi_{D,1}(x), \dots, \varphi_{D,k}(x)$ and $\varphi_{R,t}(\bar{x})$ for $R \in L$ and $t \in [k]^r$, where r is the arity of R, and x in $\varphi_{D,i}(x)$ is an individual variable, and $\bar{x}$ in $\varphi_{R,t}(\bar{x})$ is a tuple of individual variables of length r. If Ψ is such an MSO-definition and M is a structure of the vocabulary of the formulas in Ψ, then the structure defined by Ψ on M is the L-structure N with $D(N) = \{(a, i) \in D(M) \times [k] : M \models \varphi_{D,i}(a)\}$ and $R(N) = \{((a_1, i_1), \dots, (a_r, i_r)) \in (D(M) \times [k])^r : M \models \varphi_{R,(i_1,\dots,i_r)}(a_1, \dots, a_r)\}$. An MSO-definition of order k with parameters is defined analogously, with each formula carrying additional parameter variables $\bar{z}$ and $\bar{Z}$, and an additional formula $\pi(\bar{z}, \bar{Z})$ to tell if a choice of parameters $\bar{b}, \bar{B}$ is *good*. We say that Ψ takes the structure M as input and produces the structure N as output, for the good choice of parameters $\bar{b}$ and $\bar{B}$, if the defining formulas produce N when $\bar{z}$ and $\bar{Z}$ are replaced by the parameters. If there is as least one good choice of parameters in M and the same structure N is produced under all good choices of parameters, then we omit any reference to them and say that Ψ takes M as input and produces N as output.

Finally, MSO-interpretations extend MSO-definitions to allow *factor structures*. Concretely, an MSO-interpretation (without parameters, of order 1) includes an additional *equality-defining* formula $\varphi_\equiv(x, y)$, that is required to define an equivalence relation on the domain defined by $\varphi_D(x)$ that is a congruence of the relations defined by the $\varphi_R(\bar{x})$'s. On a structure M as input, the MSO-interpretation produces the structure whose domain is the set of equivalence classes of the equivalence relation $\equiv$ defined by $\varphi_\equiv(x, y)$ on the domain defined by $\varphi_D(x)$, and whose relations are the relations that are defined by the $\varphi_R(\bar{x})$'s factored by $\equiv$. MSO-interpretations with parameters and of order $k > 1$ are defined analogously.

The composition of two MSO-interpretations is defined in the obvious way and is again an MSO-interpretation. As an example of MSO-interpretation we state the following easily derived consequence of Theorem 4.7 in [3].

▶ **Theorem 3.1.** *There is an MSO-definition Ψ that, for every graph G given as input and every 3-connected component C of G, there is a good choice of parameters for Ψ on which it produces C.*

Logic Representation of Embedding Schemes. Embedding schemes will be represented as graphs expanded by two relations that represent the cyclic orderings around the vertices. Concretely, the vocabulary has two unary relation symbols V and E for vertices and edges, one binary relation symbol I for the incidence relation between vertices and edges and one ternary relation symbol R for cyclic orderings. The predicate $R(v, e_1, e_2)$ holds if e_1 and e_2 are edges that are incident to vertex v, and e_1 is the immediate predecessor of e_2 in the cyclic ordering of the edges that are incident to v.

This encoding was used by Courcelle [4] to show that there is an MSO-definition that takes a 3-connected planar graph as input and produces an embedding in the plane as output, which is unique by Whitney's Theorem.

▶ **Theorem 3.2** ([4]). *There is an MSO-definition that, given a 3-connected planar graph G as input, produces an embedding scheme as output.*

It will be convenient to extend the embedding schemes to include the faces of the embedding. Accordingly, extended embedding schemes will include a set F of faces in their domain and store the incidence relation between edges and faces through the incidence relation I.

We aim at an MSO-interpretation that given an embedding scheme produces its extension with faces. A face is determined by one of its bounding edges together with an orientation. Specifying the orientation directly by specifying one of its endpoints would take us beyond the syntax of MSO-interpretations. Instead of that, we use a proper k-coloring of the graph with a small number k of colors and specify the orientation of the edge by specifying the colors of the left and right endpoints. The k-coloring is provided through k many parameter set variables, and we choose k large enough so that any graph in the surface under consideration has chromatic number at most k. It is well-known that the chromatic number of all graphs embeddable in a fixed surface is bounded (see [10]). In the case of planar graphs $k = 5$ suffices (and even $k = 4$ does).

▶ **Lemma 3.3.** *For every $k \geq 2$, there is an MSO-interpretation that, given an embedding scheme for a graph that is k-colorable, produces its extension by faces as an extended embedding scheme.*

Embeddings by Reduction to Planar Case. Our next goal is to prove the analogue of Theorem 3.2 for higher genus surfaces. We proceed by reduction to the planar case via Theorem 2.8. Let S be an orientable surface of genus g. For simplicity we start with the orientable case.

We start with two MSO-interpretations that implement the operations of cutting a graph along a potential system of planarizing cycles, and its reverse operation of gluing along it; cf. Definition 2.7 and the discussion immediately following it. Systems are represented most simply by a sequence of $3k$ set variables.

▶ **Lemma 3.4.** *There is an MSO-interpretation Ξ that, given a graph G and potential system of planarizing cycles $\mathcal{S}$, produces the graph G′ obtained from G by cutting along $\mathcal{S}$. Conversely, there is an MSO-interpretation Ξ' that, given a graph G, a potential system of planarizing cycles $\mathcal{S}$, and a plane embedding scheme Π' for the output of Ξ on G and $\mathcal{S}$, produces the graph G″ and the embedding scheme Π'' that is obtained from G′ and Π' by gluing along $\mathcal{S}$.*

With these objects at hand we are ready to prove the analogue of Theorem 3.2.

▶ **Theorem 3.5.** *Let S be a surface of genus g. There is an MSO-interpretation that, given a 3-connected graph G that has an embedding in S of finite face-width at least $8(d+1)(2^g - 1)$ (which must be unique), produces such an embedding Π.*

We need to show that the conditions of Theorem 2.8 can be defined in MSO.

Defining Grids in High Face-Width Embeddings. Our final goal of this section is to develop an MSO-interpretation γ which defines large grids in 3-connected Π-embedded graphs.

▶ **Theorem 3.6.** *There is an MSO-interpretation γ that, given an extended embedding scheme for a graph G of finite face-width $k \geq 3$, produces a grid of order $\lfloor \frac{k}{4} \rfloor$.*

To define γ, we need to show that the conditions of Theorem 2.12 can be defined in MSO.

4 Size and face-width for random graphs

Fix a surface S of genus $g \geq 0$. It was shown in [2] that, a.a.s., a random graph G from $\mathcal{G}_S$ has genus g, and in particular it is not planar if $g > 0$. Moreover, a.a.s., G has a unique non-planar connected component, as well as a unique non-planar 2-connected and 3-connected components. Moreover, these components are all *giant*, i.e., of size linear in the number of vertices. Indeed the probability distributions of their sizes is well-understood. In the following, if f_n and g_n are sequences of positive real numbers, we use the notation $f_n \sim g_n$ to mean that $\lim_{n\to\infty} f_n/g_n = 1$.

Sizes of the components.   Let L_n denote the size (i.e., number of vertices) of the largest connected component in a random n-vertex graph G from $\mathcal{G}_S$, and let $M_n = n - L_n$. Theorem 5.3 in [2] determines the distribution of M_n, and hence of L_n: for every fixed integer $k \geq 0$ we have $\Pr[M_n = k] \sim p \cdot g_k \frac{\gamma^{-k}}{k!}$ where p, g_k and γ are constants that do not depend on n, nor on the surface S. Moreover there exists constants $a > 0$ and $b > 0$ such that $\mathrm{E}(M_n) \sim a$ and $\mathrm{Var}(M_n) \sim b$. In particular, by Chebyshev's inequality this means that, for any $a(n)$ that grows to infinity however slowly, we have $M_n \leq a(n)$ a.a.s., and hence $L_n \geq n - a(n)$ a.a.s. Theorem 5.4 and 5.5 also in [2] determine the distributions of the sizes of the largest 2-connected and 3-connected components, but only for random *connected* graphs from $\mathcal{G}_S$. However, by composing the results it is still possible to determine the sizes for random arbitrary graphs from $\mathcal{G}_S$, as we do next.

A sequence of integer random variables $X_0, X_1, \ldots$ is said to admit a local limit law of the Airy type with parameters α and c if for every real finite interval $[a, b]$ it holds that $\Pr\left[X_n = \lfloor \alpha n + xn^{2/3} \rfloor\right] \sim n^{-2/3} cg(cx)$ uniformly for every $x \in [a, b]$, where $g(x) = 2e^{-2x^3/3}(x\mathrm{Ai}(x^2) - \mathrm{Ai}'(x^2))$ and $\mathrm{Ai}(x)$ is the Airy function. Here, uniformly for every $x \in [a, b]$ means that for every positive real $\varepsilon > 0$ and every large enough n the ratio is ε-close to 1 simultanesouly for all $x \in [a, b]$.

▶ **Theorem 4.1.** *Let X_n and Y_n denote the sizes of the largest 2-connected and 3-connected components of a random n-vertex graph in $\mathcal{G}_S$. Then X_n and Y_n admit local limit laws of the Airy type (with different parameters). Moreover, a.a.s., the largest connected, 2-connected and 3-connected components are unique and have maximal possible genus, and all other connected, 2-connected and 3-connected components are planar.*

Face-width of the components.   For this section we assume that the genus of S is $g > 0$. Our goal is to show that the face-width of the largest 3-connected component of the random graph grows logarithmically. We will need the following facts:

Almost all 3-connected maps on S with m edges have face-width greater than $\delta \log m$ for some constant $\delta > 0$. This is proved in [1] for rooted maps. Since almost all 3-connected maps have no non-trivial automorphisms [12] this is also true for unrooted maps. Moreover, the largest 3-connected component of a random graph G from $\mathcal{G}_S$ is unique and non-planar by Theorem 4.1 a.a.s., and has face-width greater than any fixed constant also a.a.s. [2]. As a consequence it has a unique embedding in S, and it can be considered as an unrooted map [13].

▶ **Theorem 4.2.** *Let S be a surface of genus $g > 0$ and let F_n denote the face-width of the largest 3-connected component of a random n-vertex graph in $\mathcal{G}_S$. Then there exists γ such that $F_n \geq \gamma \log n$ asymptotically almost surely.*

Proof. Let G denote a random n-vertex graph in $\mathcal{G}_S$. Let L and T denote the largest 2-connected and 3-connected components of G. We start by arguing that, conditioned on the number of edges of L and T, the distribution of T is uniform over the 3-connected graphs in $\mathcal{G}_S$ with its number of edges. Indeed, L is obtained from T by (possibly) replacing each edge of T with a 2-connected graph, so each T with m edges gives rise to the same number L with k edges. For $m \geq k$, let $B_{m,k}$ denote the event that $e(T) = m$ and $e(L) = k$. By Theorem 4.1, the sizes of T and L are at least cn a.a.s., for some constant $c > 0$. Since they are at least 2-connected, also $e(T) \geq cn$ and $e(L) \geq cn$ a.a.s.. Now we combine the previous paragraph with the one just before the theorem: In the distribution conditioned on $B_{m,k}$, the 3-connected component T can be considered as a random m-edge 3-connected map embedded in S, and its facewidth is at least $\delta \log m$ if m is large enough. We conclude that the facewidth of T is indeed at least $\delta \log(cn)$ a.a.s. Precisely, if A denotes the event that T has facewidth at least $\delta \log(cn)$, then $\Pr[\, A \,] \geq \sum_{m \geq k \geq cn} \Pr[\, A \mid B_{m,k} \,] \Pr[\, B_{m,k} \,]$. For fixed $\varepsilon > 0$, if n is large enough, then $\Pr[\, A \mid B_{m,k} \,] \geq 1 - \varepsilon$ for any $m \geq k \geq cn$. It follows that, if n is large enough, then $\Pr[\, A \,] \geq (1 - \varepsilon) \sum_{m \geq k \geq cn} \Pr[\, B_{m,k} \,] \geq (1 - \varepsilon)^2$. Since $\varepsilon > 0$ was arbitrary, the claim is proved by choosing any $\gamma > 0$ smaller than δ. ◀

5 Limiting probabilities of MSO-sentences

In this section we put everything together. Let S be a surface of genus $g > 0$. The results so far show that a random n-vertex graph in $\mathcal{G}_S$ will have facewidth $\Omega(\log n)$ with high probability, and that on such graphs an $m \times m$ grid is MSO-definable, for some $m \geq \log \log n$. More precisely:

▶ **Theorem 5.1.** *Let S be a surface other than the sphere. There is an MSO-interpretation that, on a given n-vertex graph G from $\mathcal{G}_S$, produces an $m \times m$ grid for some $m \geq \log \log n$ for every and at least one good choice of parameters, asymptotically almost surely when G is a random n-vertex graph in $\mathcal{G}_S$.*

We use this to build MSO-sentences with non-converging probabilities. We use it also for proving the undecidability of the problem of determining the asymptotic probabilities of MSO-sentences. For building MSO-sentences whose probabilities converge to any given rational number in the interval $[0, 1]$ we use Theorem 4.1 from the previous section.

MSO-sentences with non-converging probabilities. For every base $b \geq 2$ and every natural number n, define $\mathrm{tow}_b(n)$ recursively by $\mathrm{tow}_b(0) = 1$ and $\mathrm{tow}_b(i + 1) = b^{\mathrm{tow}_b(i)}$. For every real $x \geq 0$, let $\log_b^*(x)$ denote the smallest integer k such that $\mathrm{tow}_b(i) \geq x$, and let $\log_b^\dagger(x)$ denote the smallest integer k such that $\sum_{i=0}^{k} \mathrm{tow}_b(i) \geq x$. By induction on k one proves that $\sum_{i=0}^{k} \mathrm{tow}_b(i) \leq 2\mathrm{tow}_b(k)$, and hence $\log_b^*(x/2) \leq \log_b^\dagger(x) \leq \log_b^*(x)$ for every $b \geq 2$ and every real $x \geq 0$. Both $\log_b^*(n)$ and $\log_b^\dagger(n)$ are monotone non-decreasing functions of n that have all natural numbers in their range. When the base is 2 we omit b from the notation.

The source of divergence in our example is the following easily verified arithmetic fact:

▶ **Lemma 5.2.** *If $m_1, m_2, \ldots$ is an integer sequence such that $\log \log(n) \leq m_n \leq n$ for every large enough n, then the sequence given by $\log^\dagger(m_n) \bmod 8$ is not eventually always in $\{0, 1, 2, 3\}$ and not eventually always in $\{4, 5, 6, 7\}$.*

For the next technical lemma it will be more convenient to move, temporarily, to the vocabulary of directed grids. For every $i \in \{0, \ldots, 7\}$, we want an MSO-sentence $\mathtt{strange}_i$ that holds in the $n \times n$ directed grid if and only if $\log^\dagger(n)$ is congruent to i mod 8. We use the notation $G_{n \times n}^{\mathrm{d}}$ to denote the $n \times n$ directed grid.
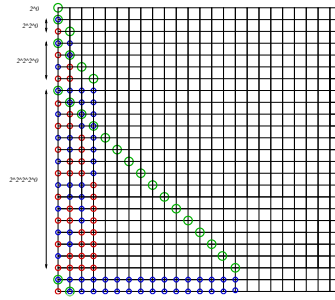
■ **Figure 5.1** The red-green-blue pattern for the proof of Lemma 5.3.

▶ **Lemma 5.3.** *For every $i \in \{0, \ldots, 7\}$ there exists an MSO-sentence $\mathtt{strange}_i$ in the vocabulary of directed grids such that, for every natural number $n \geq 1$, the sentence $\mathtt{strange}_i$ is true in $G_{n \times n}^{\mathrm{d}}$ if and only if $\log^{\dagger}(n)$ is congruent to i mod 8.*

Proof. The sentence uses three existentially quantified monadic second-order variables R, G, and B, for red, green and blue. First it verifies that the colors satisfy the pattern of Figure 5.1. Once this is verified, the sentence states that the number of green vertices in the leftmost column is congruent to $i + 1$ mod 8. The number of green vertices in the leftmost column is the smallest k such that $\sum_{i=0}^{k-1} \mathrm{tow}(i) \geq n$; i.e. $k = \log^{\dagger}(n) + 1$, so the statement states that $\log^{\dagger}(n) + 1 \equiv i + 1$ mod 8, which is the same as $\log^{\dagger}(n) \equiv i$ mod 8.   ◄

We need to find a way of defining directed grids in the $\{\mathtt{right}, \mathtt{down}\}$-vocabulary from undirected ones. One way to do this via MSO-interpretations can be extracted from Section 5.2.3 in [5].

▶ **Theorem 5.4.** *Let S be a surface other than the sphere. There is an MSO-sentence whose asymptotic probability on $\mathcal{G}_S$ does not converge.*

Proof. Let Θ be the composition of Ψ from Theorem 5.1 with interpretation that produces directed grids from undirected ones. For every $T \subseteq \{0, \ldots, 7\}$, let φ_T say that there is a good choice of the parameters for Θ that make it define a directed square grid on which the disjunction $\bigvee_{i \in T} \mathtt{strange}_i$ holds. If $\varphi_{\{0,1,2,3\}}$ has probability 0, then $\varphi_{\{4,5,6,7\}}$ has probability 1 by Lemma 5.3. Let $T \in \{\{0,1,2,3\}, \{4,5,6,7\}\}$ be such that φ_T does not have asymptotic probability 0.

We claim that the asymptotic probability of φ_T does not converge. Otherwise, it converges to a positive real, and a positive fraction of the n-vertex graphs in $\mathcal{G}_S$ satisfy φ_T. Therefore, for n there is a least one n-vertex graph in $\mathcal{G}_S$ for which there is a good choice of parameters that makes Ψ define a directed $m \times m$ grid with $\log^{\dagger}(m)$ mod 8 in T and such that at the same time $m \geq \log \log n$. In other words, we find a sequence m_n that contradicts Lemma 5.2.   ◄

Undecidability of the decision problem.   We proceed by reduction from the halting problem for Turing machines, which is of course undecidable. For every 1-tape Turing machine M we want an MSO sentence $\mathtt{halts}_M$ of the vocabulary of directed grids, that holds in the $n \times n$ directed grid if and only if the computation of M on the empty input halts in time at most n using space at most n. This construction is standard.

▶ **Lemma 5.5.** *For every two-sided 1-tape Turing machine M there exists an MSO-sentence $\mathtt{halts}_M$ of the vocabulary of directed grids such that, for every natural number $n \geq 1$, the sentence halts_M is true in $G_{n \times n}^{\mathrm{d}}$ if and only if the computation of M on the empty input halts in time at most n using space at most n.*

We combine the non-converging formula of Theorem 5.4 with Lemma 5.5.

▶ **Theorem 5.6.** *Let S be a surface other than the sphere. The problem of determining whether a given MSO-sentence has a converging asymptotic probability on $\mathcal{G}_S$ is undecidable.*

Proof. Let φ be the MSO-sentence with non-converging asymptotic probability from Theorem 5.4. As in the proof of Theorem 5.4, let Ψ be the MSO-interpretation in Theorem 5.1, and let Θ be the composition of Ψ with the MSO-interpretation that produces a directed grid from an undirected one. For every two-sided 1-tape Turing machine M, let ψ_M be the sentence that says that there is a good choice of parameters for Θ that makes it define a square grid, and the sentence $\texttt{halts}_M$ holds on some principal square subgrid of this grid. It is easy to see that ψ_M has asymptotic probability one if M halts, and probability zero otherwise. This $\varphi \wedge \psi_M$ converges if and only if M halts.   ◀

All rationals in [0,1] as limiting probabilities.   Our goal in this section is to construct an MSO sentence whose asymptotic probability over $\mathcal{G}_S$, for any fixed surface S other than the sphere, converges to any given rational number in the interval $[0, 1]$. The idea of the construction is the following. Let us say that we want to achieve the rational p/q as limiting probability. Suppose that we succeed to write a sentence that says that the unique non-planar 3-connected component of the random n-vertex graph has size that is congruent to some $a \in \{0, \ldots, p-1\}$ mod q. If we do, then by Theorem 4.1 the probability that this sentence holds on the random n-vertex graph is the probability that an integer random variable that admits a local limit law of the Airy type is congruent to some $a \in \{0, \ldots, p-1\}$ mod q. It turns out that this probability approaches p/q as n approaches infinity:

▶ **Lemma 5.7.** *Let $X_0, X_1, \ldots$ be a sequence of integer random variables that admits a local limit law of the Airy type with parameters α and c. Then for every integer $q \geq 1$ and every $a \in \{0, \ldots, q-1\}$ it holds that $\lim_{n \to \infty} \Pr\left[X_n \equiv a \ (\mathrm{mod} \ q)\right] = 1/q$.*

For saying that the unique non-planar 3-connected component has a size that is congruent to a mod q we use the fact every 3-connected graph of bounded Euler characteristic $\chi \leq 0$ has a spanning tree of degree at most $\lceil(8 - 2\chi)/3\rceil$ [7, 11]. Thus, the unique non-planar 3-connected component, which is embeddable in the surface S, has such a spanning tree, which can be guessed in MSO_2. Once available, the spanning tree of bounded degree can be used to define a linear order on the vertices of the 3-connected component, and MSO over the linear order can say that its length is congruent to a mod q. Taking the disjunction over all $a \in \{0, \ldots, p-1\}$, the asymptotic probability of the resulting sentence will be p/q.

▶ **Theorem 5.8.** *Let S be a surface other than the sphere. For every rational number $r \in [0, 1]$, there exists an MSO sentence whose asymptotic probability on $\mathcal{G}_S$ converges to r.*

───── **References** ─────

1   E. A. Bender, Z. Gao, L. B. Richmond, and N. Wormald. Asymptotic properties of rooted 3-connected maps on surfaces. *J. Austral. Math. Soc., Series A*, 60:31–41, 1996.

2   G. Chapuy, E. Fusy, O. Giménez, B. Mohar, and M. Noy. Asymptotic enumeration and limit laws for graphs of fixed genus. *J. Combin. Theory Ser. A*, 118:748–777, 2011.

3   B. Courcelle. The monadic second-order logic of graphs XI: Hierarchical decompositions of connected graphs. *Theoretical Computer Science*, 224:35–58, 1999.

4   B. Courcelle. The monadic second-order logic of graphs xii: planar graphs and planar maps. *Theoretical Computer Science (TCS)*, 237:1–32, 2000.

5   B. Courcelle and J. Engelfriet. *Graph Structure and Monadic Second-Order Logic.* Cambridge University Press, 2012.

6   R. Diestel. *Graph Theory.* Springer-Verlag, 4th edition, 2010.

7   M.N. Ellingham. Spanning paths, cycles and walks for graphs on surfaces. In *Congressus Numerantium*, volume 115, pages 55–90, 1996.

8   O. Giménez, M. Noy, and J. Rué. Graph classes with given 3-connected components: asymptotic enumeration and random graphs. *Random Structures Algorithms*, 42(4):438–479, 2013.

9   P. Heinig, T. Müller, M. Noy, and A. Taraz. Logical limit laws for minor-closed classes of graphs. *J. Combin. Theory Ser. B*, to appear.

10   B. Mohar and C. Thomassen. *Graphs on Surfaces.* Johns Hopkins University Press, 2001.

11   K. Ota and K. Ozeki. Spanning trees in 3-connected $K_{3,t}$-minor-free graphs. *J. Combin. Theory Ser. B*, 102(5):1179–1188, 2012.

12   L. B. Richmond and N. C. Wormald. Almost all maps are asymmetric. *J. Combin. Theory Ser. B*, 63(1):1–7, 1995.

13   N. Robertson and R. Vitray. Representativity of surface embeddings. In *Paths, flows, and VLSI-layout (Bonn, 1988)*, volume 9 of *Algorithms Combin.*, pages 293–328. Springer, Berlin, 1990.