Chapter 1

Local convergence of graphs

In this chapter we lay down the basics of graph theory and present the local topology on the space of pointed graphs. This Polish topology will enable us to state proper convergence results for graphs of growing size towards infinite random limits.

1.1 Graphs, examples, questions

1.1.1 Graphs

In these notes, a **graph** (sometimes also called a non-oriented multi-graph) is a pair $\mathsf{g} = (\mathrm{V}(\mathsf{g}), \mathrm{E}(\mathsf{g}))$, where $V = \mathrm{V}(\mathsf{g})$ is the set of vertices of g and $E = \mathrm{E}(\mathsf{g})$ is the set of edges of g which is a multiset over $\{V\}^2$, the set of unordered pairs of elements of $\mathrm{V}(\mathsf{g})$. The graph is **simple** if there are no multiple edges or loops. If $x, y \in V$ and $\{x, y\} \in E$

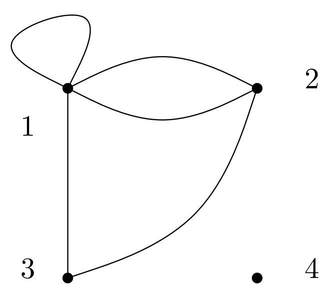


Figure 1.1: An example of a graph $\mathsf{g} = (V, E)$ with vertex set $V = \{1, 2, 3, 4\}$ and edge set $E = \{\!\{\{1, 1\}, \{1, 2\}, \{1, 2\}, \{1, 3\}, \{3, 2\}\}\!\}$. The vertex degrees of $1, 2, 3, 4$ are respectively $5, 3, 2, 0$.

we say that x and y are neighbors and that x and y are adjacent to the edge $\{x, y\}$. The **degree** of a vertex $x \in V$, denoted by $\deg(x)$, is the number of half-edges adjacent to x; otherwise said, it is the number of edges adjacent to x, where loops are counted twice. A graph g has bounded degree if $\sup_{x\in \mathrm{V}(\mathsf{g})} \deg(x) < \infty$. The graph distance on g is denoted by $\mathrm{d}_{\mathrm{gr}}^{\mathsf{g}}$, or d_{gr} when there is no ambiguity, and is defined by

$$\mathrm{d}_{\mathrm{gr}}(x,y) = \text{minimal number of edges to cross in order to go from } x \text{ to } y.$$

By convention, we set $\mathrm{d}_{\mathrm{gr}}(x,y) = \infty$ if there is no path linking x to y in g. The equivalence classes for the relation $\sim$, where $x \sim y \iff \mathrm{d}_{\mathrm{gr}}(x,y) < \infty$, are the connected components of g. We say that g is connected if it has only one connected component.

Proposition 1. *For a connected graph* $\mathsf{g} = (V,E)$ *on* n *vertices, we must have* $\#E \geq n-1$. *If* $\#E = n-1$ *then* g *is a tree, meaning that is has no non-trivial cycle.*

Proof. We can assume that the vertex set of g is $\{1,2,3,...,n\}$. We start by considering vertex 1; Since g is connected there exists an edge adjacent to 1 of the form $\{1,i_1\}$. If $i_1 = 1$ then this edge is a loop, otherwise $i_1 \neq 1$. We then mark this edge as explored and pick a new edge adjacent to either 1 or i_1. Iteratively after having marked k edges, we have explored a part of the connected component of 1 which has at most $k+1$ vertices. Since g is connected it follows that $\#E \geq n-1$. In case of equality this means that during the exploration process we have never found an edge linking two vertices already explored, in other words no non trivial cycle has been created and thus g is a tree. ◻

Graph equivalence. If g and g' are two graphs we say that g and g' are equivalent if they represent the same graph up to renaming the vertex set. Formally this means that there exists a bijection $\phi : \mathrm{V}(\mathsf{g}) \to \mathrm{V}(\mathsf{g}')$ which maps the multiset $\mathrm{E}(\mathsf{g})$ to $\mathrm{E}(\mathsf{g}')$: such a function is called a graph homomorphism (automorphism if $\mathsf{g} = \mathsf{g}'$). Geometrically, two equivalent graphs are the same and we shall by abuse of notion often speak use the term graph and the notation g for its equivalence class. Although the set of equivalence classes of finite graphs is a discrete space, the space of equivalence classes of infinite graphs is monstrous even in the case of countable graphs, see [67]. The salvation in this course will come from the introduction of an **origin** for the graph g, which will either be a distinguished vertex or a distinguished oriented edge.

1.1.2 Regular graphs

In this section we introduce several classes of "regular" graphs in order to motivate the forthcoming definitions of stationary or unimodular random graphs, which in a sense are generalizations of the following notions.

Definition 1 (Cayley graph)**.** *A Cayley graph encodes the structure of a group. Given a (countable) group* Gr *and a finite symmetric generating set* $S = \{s_1, s_1^{-1}, s_2, s_2^{-1}, ..., s_k, s_k^{-1}\}$ *of* Gr*, we form the simple graph whose vertices are the elements of the group* Gr *and an edge links* x *to* y *if there exists* $s \in S$ *such that* $x = sy$ *in* Gr*.*

In particular, a Cayley graph depends not only upon the group Gr but also on the symmetric generating set we have chosen. For example, here are two Cayley graphs for the group $(\mathbb{Z}, +)$, one corresponding to the symmetric generating set $\{-1, +1\}$ and the other to $\{-3, -2, 2, 3\}$:

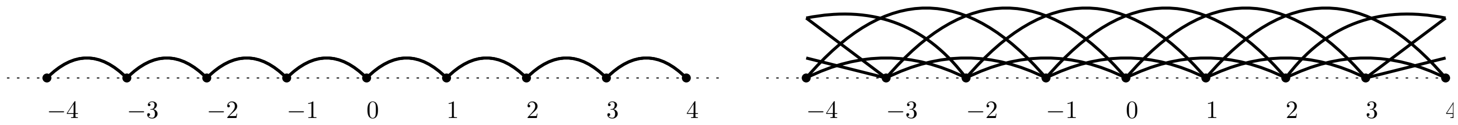


Figure 1.2: Two Cayley graphs of the same group $(\mathbb{Z}, +)$

In particular, Cayley graphs are connected and "regular" in the sense that they look the same when seen from any vertex. This is formalized in the following notion:

Definition 2 (Transitive graphs)**.** *A graph* $\mathsf{g} = (V, E)$ *is (vertex-)transitive if for any* $u, v \in V$ *there is a graph automorphism that takes* u *to* v *(in pedantic terms, the automorphism group of* g *acts transitively on its vertices).*

Clearly, any Cayley graph is vertex-transitive, but the converse is false as we will see later. In fact, as it will turn out, Cayley graphs possess, on top of transitivity, an additional "inverse" property which is not always present in infinite vertex-transitive graphs. A transitive graph is in particular **d-regular** for some $d \in \{1, 2, 3...\}$, i.e., the vertex degrees are constant and equal to d. In the sequel we shall write $\mathbb{E}^d$ for the d-dimensional Euclidean lattice obtained as the Cayley graph of $\mathbb{Z}^d$ with the standard basis $\pm(0, ...0, 1, 0, ..., 0)$ of generators. We also write $\mathbb{E}^d_+$ for the previous graph restricted to $\mathbb{Z}^d_+$. We write $\mathbb{T}_d$ for the d-regular infinite tree.

Exercise 1. Show that the Petersen graph is transitive but is not a Cayley graph:

We could go on and continue weakening the notion of regularity by introducing quasi-transitive graphs (where the action of the automorphism group has only finitely many orbits)... However, since we will be later interested in random graphs, any deterministic notion of regularity is doomed. The reader should however keep in mind the last two notions since they will play a major role in these notes.

1.1.3 Examples of random graphs

We informally introduce several models of random graphs that we shall discuss in these pages:

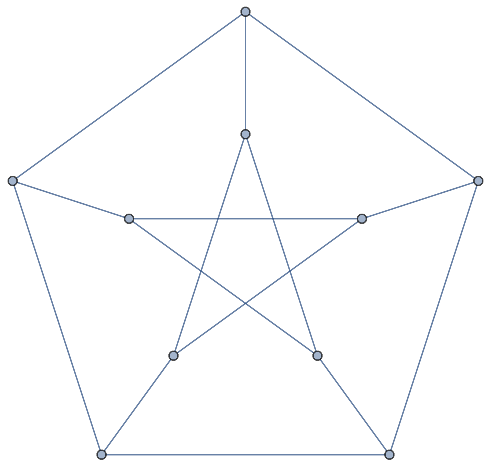


Figure 1.3: The Petersen graph

Percolation. Given a finite or countable graph g, we call Bernoulli bond (resp. site) percolation on g with parameter $p \in (0,1)$ the random graph obtained from g by keeping independently each edge (resp. vertex) with probability p and erasing it with probability $1-p$.

- When the underlying graph is the complete graph on n vertices (there is an edge between any pair of distinct vertices, so there are $n(n-1)/2$ edges in total) and we perform Bernoulli bond percolation, we speak of the **Erdős–Rényi random graph**, denoted by $G(n,p)$. Introduced in 1959, it is the most famous and most studied model of random graph. This model is referred to as "mean-field" because the geometry of the underlying lattice is trivial in the sense that any pair of vertices are neighbors of each other.

- When the underlying lattice is the classical lattice $\mathbb{E}^d$ for $d \geq 2$, it corresponds to the well-known model of Bernoulli bond or site percolation on regular lattices, first studied by Broadbent and Hammersley in 1957 [31]. The main question is then the existence of an infinite cluster. The general theory shows that there is a critical percolation threshold $p_c(\mathbb{E}^d) \in (0,1)$ such that for $p < p_c(\mathbb{E}^d)$ there is no infinite cluster and for $p > p_c(\mathbb{E}^d)$ there is a unique infinite cluster.

The percolation model has many variants (inhomogeneous percolation, stochastic block models) that may model real networks with more accuracy.

Galton–Watson tree. Given a probability distribution $\mathbf{p} = (p_k)_{k\geq 0}$ on $\{0,1,2,...\}$, a Galton–Watson tree with offspring distribution $\mathbf{p}$ is informally the genealogical tree obtained by starting from a single ancestor particle and letting each particle reproduce independently of the others according to the offspring distribution $\mathbf{p}$. This model of a random tree is also very well known and very well understood, see Section 3.1.1. It is one of the key characters in the theory of random graphs.

Uniform graph in a combinatorial class. Let $\mathscr{C}_n$ be a combinatorial class of graphs with finite size. We can for example take $\mathscr{C}_n$ to be the set of all graphs with n edges, the set of all graphs with n edges and $\alpha(n)$ vertices, the set of all trees with n vertices, the set of all planar graphs with n edges... Provided that $\#\mathscr{C}_n$ is finite (which is the case in the above examples) then we can consider $C_n \in \mathscr{C}_n$, a random variable uniformly distributed over $\mathscr{C}_n$.

Dynamical random graphs. We can also consider a sequence of growing random graphs that are built recursively. The (very interesting) prototype is the Barabási–Albert model [14], which was introduced to model the degree distribution of many real networks such as the internet. The initial graph G_1 is just a single vertex. Then inductively for $n \geq 1$, we choose in G_n a vertex proportionally to its degree (the number of adjacent edges) and we attach a new leaf to this vertex. Hence, the vertices with higher degree are more likely to get chosen: the rich get richer.

Configuration model. The configuration model is designed so that the degree sequence of the vertices is fixed a priori. More precisely, let $\mathbf{d} = (d_1, ..., d_n)$ be a sequence of integers such that $D = \sum d_i$ is even. We will then consider a random (multi-)graph on n vertices $\{1, 2, ..., n\}$ such that the vertex i has degree d_i. One way to sample such a graph is to start with those n vertices to which we attach d_i half-edges (also called stubs), which are labeled from 1 to D. We then pair those half-edges in the most natural way: we match the half-edge number 1 with a uniformly chosen half-edge among those numbered $2, 3, ..., D$. We then merge the two stubs involved and create a true edge. We iterate the procedure with the remaining stubs and denote by $G(\mathbf{d})$ the random graph obtained. This graph may not be simple (it may contain multiple edges or loops).

1.1.4 In search of a limit model

In most of the above examples, the random graphs that are considered have finite size n. It would be desirable to define a limiting infinite model when $n \to \infty$ that captures ”intensive” variables. Indeed, it is common in mathematics that asymptotic questions on models of growing size could often be resolved by a direct analysis of an appropriate infinite limiting model (which usually has nicer properties) provided of course that the desirable observables are continuous. Let us give three examples of such questions: Consider (G_n)

is a sequence of (random) graphs whose size tends to infinity. One may be interested in investigating:

Degree distribution. the degree distribution, i.e., the random vector

$$\mathbf{p}_n(k) = \frac{\#\{u \in \mathrm{V}(G_n) : \deg(u) = k\}}{\#\mathrm{V}(G_n)}, \quad \text{as } n \to \infty.$$

Percolation giant density. Consider a Bernoulli bond percolation on G_n with parameter p and consider $\theta_n(p)$ the averaged density of the largest cluster in the percolated graph.

Spanning trees count. A spanning tree of a (connected) graph g is a connected subgraph of g without cycles which spans all the vertices of g. In the growing sequence (G_n), one may would like to understand the asymptotic number of spanning trees

$$\frac{\log \#\mathrm{SpanTrees}(G_n)}{|G_n|}, \quad \text{as } n \to \infty.$$

Matching number. A matching on g is a subset of mutually non-adjacent edges on g. We denote by $\nu(\mathrm{g})$ the largest size of a matching on g. Again if (G_n) is a sequence of (random) graphs whose size tends to infinity, one may would like to understand the asymptotic

$$\frac{\nu(G_n)}{|G_n|}, \quad \text{as } n \to \infty.$$

All these asymptotic enumeration problems can indeed be answered [51, 2] by looking at an appropriate limit of the sequence (G_n), the so-called Benjamini-Schramm local limit. The goal of this course is to describe this notion of limit, which is well-suited to the analysis of so-called "dilute" random graphs where the average number of edges per vertex (or mean degree) remains typically bounded.

From now on, unless explicitly mentioned, all the graphs g considered are

- countable, i.e., $\mathrm{V}(\mathrm{g})$ is countable,
- locally finite, i.e., $\deg(x) < \infty$ for all $x \in \mathrm{V}(\mathrm{g})$,
- connected.

1.2 Local convergence topology

Let us now develop the concept of local topology in a quite abstract setup because we shall need to apply it to several different cases: pointed graphs, bi-pointed graphs, rooted graphs, graph with a distinguished path, with or without an orientation, with a labeling...

1.2.1 Local topology

Imagine that we are given a space $\mathcal{E}$ such that for any $x \in \mathcal{E}$ and for any $r \in \{0, 1, 2, 3, ...\}$, there is a notion of **restriction of radius** r of x, that we denote by

$$[x]_r \in \mathcal{E}.$$

We also denote by $[\mathcal{E}]_r = \{[x]_r : x \in \mathcal{E}\}$. Suppose that we have a distance δ on $\mathcal{E}$, and that the following assumptions are satisfied:

- **Compatibility:** different restrictions of radius r are compatible in the sense that

$$[[x]_r]_{r'} = [x]_{r'}, \qquad \text{for any } r \geq r' \geq 0,$$

- **Continuity:** for any $r \geq r' \geq 0$, the restriction map $[\cdot]_{r'} : ([\mathcal{E}]_r, \delta) \to ([\mathcal{E}]_{r'}, \delta)$ is continuous,

- **Polishness of $[\mathcal{E}]_r$:** for any $r \geq 0$, the metric space $([\mathcal{E}]_r, \delta)$ is separable and complete,

- **Projective limit:** for any coherent sequence $x_0, x_1, ... \in \mathcal{E}$, i.e. such that $[x_r]_{r'} = x_{r'}$ for any $r \geq r' \geq 0$, there exists a unique "*infinite*" element $x \in \mathcal{E}$ such that

$$[x]_r = x_r, \qquad \text{for all } r \geq 0. \tag{1.1}$$

We denote these assumptions by $(*)$. We then endow the space $\mathcal{E}$ with the following distance $\mathrm{d_{loc}}$, called the **local distance**, which is defined as

$$\mathrm{d_{loc}}(x, y) = \sum_{r \geq 0} \min(\delta([x]_r, [y]_r), 1) \cdot 2^{-r}.$$

In other words, a sequence (x_n) of elements of $\mathcal{E}$ converges towards x for the local distance if and only if for any $r \geq 0$, the sequence $[x_n]_r$ converges to $[x]_r$ for the metric δ as $n \to \infty$. We shall use many times below that

$$\mathrm{d_{loc}}(x, [x]_r) \leq \sum_{k > r} 2^{-k} \leq 2^{-r}.$$

In this general setting, we have the following:

Theorem 2 (*The local topology is Polish*)

Under the assumptions $(*)$, the metric space $(\mathcal{E}, \mathrm{d_{loc}})$ is separable and complete. Furthermore, a subset $\mathcal{A} \subset \mathcal{E}$ is pre-compact for $\mathrm{d_{loc}}$ (meaning that its closure is compact) *if and only if*

$$\text{for every } r \geq 0, \text{ the subset } [\mathcal{A}]_r = \{[x]_r : x \in \mathcal{A}\} \text{ is pre-compact for } \delta.$$

Proof. Distance. Let us first show that $\mathrm{d}_{\mathrm{loc}}$ is a distance. The symmetry is easy, as well as the triangle inequality. The separability is also easy since we suppose that if $[x]_r = [y]_r$ for all $r \geq 0$ then we have $x = y$ by (Projective limit). Let us note for future use that for each $r \geq 0$, the restriction map $[\cdot]_r : (\mathcal{E}, \mathrm{d}_{\mathrm{loc}}) \mapsto ([\mathcal{E}]_r, \delta)$ is uniformly continuous: this follows from the inequality

$$\min(\delta([x]_r, [y]_r), 1) \leq 2^r \cdot \mathrm{d}_{\mathrm{loc}}(x, y) \quad \text{for all } x, y \in \mathcal{E}. \tag{1.2}$$

Separability. Recall that a metric space (E, d) is separable if and only if for each $\varepsilon > 0$ there exists a sequence $x_1, x_2, \ldots \in E$ such that $E = \bigcup_{i \geq 1} B_d(x_i, \varepsilon)$. Now, fix $\varepsilon > 0$, and fix $R \geq 0$ large enough so that $2^{-R} \leq \varepsilon$. For each $0 \leq r \leq R$, by the separability of $([\mathcal{E}]_r, \delta)$, there exists a sequence $x_1^r, x_2^r, \ldots \in [\mathcal{E}]_r$ such that $[\mathcal{E}]_r \subset \bigcup_{i \geq 1} B_\delta\left(x_i^r, \varepsilon\right)$. Then, consider the subsets

$$O_{i_0,\ldots,i_R} = \left\{x \in \mathcal{E} : [x]_r \in B_\delta\left(x_{i_r}^r, \varepsilon\right) \text{ for all } 0 \leq r \leq R\right\},$$

for $i_0, \ldots, i_R \geq 0$. These are open subsets for $\mathrm{d}_{\mathrm{loc}}$ (by the continuity of the restriction maps $[\cdot]_r : (\mathcal{E}, \mathrm{d}_{\mathrm{loc}}) \to ([\mathcal{E}]_r, \delta)$) which cover $\mathcal{E}$, and have diameter at most 5ε for $\mathrm{d}_{\mathrm{loc}}$: for every $x, y \in O_{i_0,\ldots,i_R}$, we have

$$\mathrm{d}_{\mathrm{loc}}(x, y) \leq \sum_{r=0}^{R} 2\varepsilon \cdot 2^{-r} + \sum_{r \geq R+1} 2^{-r} \leq 4\varepsilon + 2^{-R} \leq 5\varepsilon. \tag{1.3}$$

The result follows.

Completeness. If $(x_n)_{n \geq 1}$ is a Cauchy sequence for $\mathrm{d}_{\mathrm{loc}}$, then for each r its restriction $([x_n]_r)_{n \geq 1}$ is a Cauchy sequence for δ (this follows from (1.2)). By the completeness of $([\mathcal{E}]_r, \delta)$, we get that $([x_n]_r)_{n \geq 1}$ converges for δ to a certain element $y_r \in [\mathcal{E}]_r$. By (Continuity and Compatibility), for any $r \geq r' \geq 0$, we have

$$y_{r'} = \lim_{n \to \infty} [x_n]_{r'} = \lim_{n \to \infty} [[x_n]_r]_{r'} = [y_r]_{r'},$$

and so by (Projective limit) there exists a unique element $y \in \mathcal{E}$ such that $y_r = [y]_r$ for all $r \geq 0$. It is then immediate that $x_n \to y$ for $\mathrm{d}_{\mathrm{loc}}$ since for each $r \geq 0$, we have $[x_n]_r \to [y]_r$ for δ.

Characterization of pre-compact subsets. On the one hand, if $\mathcal{A}$ is pre-compact for $\mathrm{d}_{\mathrm{loc}}$, then for each $r \geq 0$, its restriction $[\mathcal{A}]_r$ is pre-compact for δ by the continuity of the restriction map $[\cdot]_r : (\mathcal{E}, \mathrm{d}_{\mathrm{loc}}) \to ([\mathcal{E}]_r, \delta)$. To prove the converse implication, recall that in a complete metric space (E, d), a subspace $A \subset E$ is pre-compact if and only if it is totally bounded, i.e, for each $\varepsilon > 0$, there exists finitely many elements $x_1, \dots, x_k \in A$ such that $A \subset \bigcup_{i=1}^{k} B_d(x_i, \varepsilon)$. Now, fix $\varepsilon > 0$, and fix $R \geq 0$ large enough so that $2^{-R} \leq \varepsilon$. For each $0 \leq r \leq R$, by the pre-compactness of $[\mathcal{A}]_r$ for δ, there exists finitely many elements $y_1^r, \dots, y_{k_r}^r \in [\mathcal{A}]_r$ such that $[\mathcal{A}]_r \subset \bigcup_{i=1}^{k_r} B_\delta\left(y_i^r, \varepsilon\right)$. Then, consider the subsets

$$O_{i_0,\dots,i_R} = \left\{x \in \mathcal{E} : [x]_r \in B_\delta\left(y_{i_r}^r, \varepsilon\right) \text{ for all } 0 \leq r \leq R\right\},$$

for $0 \leq i_0 \leq k_0, \dots, 0 \leq i_R \leq k_R$. These are open subsets for $\mathrm{d_{loc}}$ (by the continuity of the restriction maps $[\cdot]_r : (\mathcal{E}, \mathrm{d_{loc}}) \to ([\mathcal{E}]_r, \delta)$) which cover $\mathcal{A}$, and have diameter at most 5ε for $\mathrm{d_{loc}}$, by the same calculation as in (1.3). The result follows. $\square$

Let us give some examples:

Exercise 2. Consider the space $(\mathcal{E}, \delta) = (\mathcal{C}(\mathbb{R}_+, [-1, 1]), \|\cdot\|_\infty)$, endowed with the restriction operation:

$$[f]_r(x) = f(x)\mathbf{1}_{x \leq r} + f(x)(r + 1 - x)\mathbf{1}_{r \leq x \leq r+1}.$$

Show that $(*)$ holds and that the local topology coincides with the uniform convergence over all compact sets of $\mathbb{R}_+$ (which is then Polish).

Exercise 3. Consider the space $\mathcal{E} = [0, 1]$ endowed with the trivial distance $\delta(x, y) = \mathbf{1}_{x \neq y}$ (in particular not Polish). The restriction operation $[x]_r$ consists in truncating the binary expansion of x after r digits. Show that $(*)$ are satisfied and describe the local topology on $[0, 1]$ induced.

More generally when the space $\mathcal{E}$ is seen as a **discrete** space, it may be endowed with the trivial distance. Checking the assumptions $(*)$ then boils down to verifying the compatibility and injectivity of the restriction operation and the fact that $[\mathcal{E}]_r$ is countable for every $r \geq 0$. Then the above process enables to put a local distance on $\mathcal{E}$ making it a Polish space. In these lecture notes we will apply this construction of the local topology to the following discrete structures (see below for details):

- The space $\mathcal{G}^\bullet$ of equivalence classes of pointed graphs,
- The space $\vec{\mathcal{G}}$ of equivalence classes of rooted graphs,
- The space $\mathcal{G}^\leftrightarrow$ of equivalence classes of graphs endowed with a path,
- The space $\mathcal{M}^\bullet$ of pointed planar maps,
- The space $\vec{\mathcal{M}}$ of rooted planar maps,
- The space $\mathcal{T}$ of rooted plane trees,

We will always write $\mathrm{d}_{\mathrm{loc}}$ for the local distance and it will be clear from the context what exactly we mean. In the following exercises, we propose to re-interpret well-known topologies as local topologies:

Exercise 4 (Fell topology). Recall that the Fell topology (or hit-and-miss topology) on the set $\mathcal{F}$ of all closed subsets of $\mathbb{R}^d$ is generated by the sets of the form

$$\{F \in \mathcal{F} : F \cap K = \varnothing\} \quad \text{and} \quad \{F \in \mathcal{F} : F \cap U \neq \varnothing\},$$

for $K \subset \mathbb{R}^d$ compact or $U \subset \mathbb{R}^d$ open, see [16, Chapter 5]. For $r \geq 0$, let us denote by $\overline{B}_r$ the closed ball of radius r in $\mathbb{R}^d$ and consider the restriction operators

$$F \in \mathcal{F} \mapsto [F]_r = (F \cap \overline{B}_r) \cup \partial \overline{B}_r.$$

Show that the Fell topology can be seen as the local topology induced by the restrictions above and endowing $[\mathcal{F}]_r$ with the Hausdorff distance on closed subsets of a compact space, see [16, Chapter 3].

Exercise 5. Re-interpret [6] in the above framework.

Finally, we conclude this section by discussing a more general instance of local topology. In some cases, such as Exercise 6 below, we are not actually given a space $(\mathcal{E}, \delta)$ together with restriction maps $[\cdot]_r$ for $r \geq 0$. Rather, we are given for each $r \geq 0$ a space of restrictions $(\mathcal{E}_r, \delta_r)$, together with restriction maps $[\cdot]^r_s : \mathcal{E}_r \to \mathcal{E}_s$ for $r \geq s \geq 0$. In this case, we assume that the following holds:

- **Compatibility:** for every $r \geq s \geq t \geq 0$, we have $[\cdot]^r_t = [\cdot]^s_t \circ [\cdot]^r_s$, and $[\cdot]^r_r = \mathrm{id}_{\mathcal{E}_r}$,
- **Continuity:** for every $r \geq s \geq 0$, the restriction map $[\cdot]^r_s : (\mathcal{E}_r, \delta_r) \to (\mathcal{E}_s, \delta_s)$ is continuous,
- **Polishness of $\mathcal{E}_r$:** for each $r \geq 0$, the metric space $(\mathcal{E}_r, \delta_r)$ is separable and complete.

Although the restriction spaces need not be included in a common space, one can emulate such a space by considering their so-called projective limit, the space of coherent sequences:

$$\mathcal{E} = \left\{ x = (x_r)_{r \geq 0} \in \prod_{r \geq 0} \mathcal{E}_r : [x_r]^r_s = x_s \text{ for all } r \geq s \geq 0 \right\}.$$

In this setting, we consider the restriction maps

$$\begin{array}{rccc} [\cdot]_r : & \mathcal{E} & \longrightarrow & \mathcal{E}_r \\ & x & \longmapsto & x_r, \end{array}$$

which are compatible in the following sense: for every $r \geq s \geq 0$, we have $[\cdot]_s = [\cdot]_s^r \circ [\cdot]_r$. Finally, we define the local distance $\mathrm{d}_{\mathrm{loc}} : \mathcal{E} \times \mathcal{E} \to \mathbb{R}_+$ by

$$\mathrm{d}_{\mathrm{loc}}(x, y) = \sum_{r \geq 0} \min(\delta_r(x_r, y_r), 1) \cdot 2^{-r} \quad \text{for all } x, y \in \mathcal{E}.$$

One can then check that the analogue of Theorem 2 holds in this more general setting, namely: the metric space $(\mathcal{E}, \mathrm{d}_{\mathrm{loc}})$ is separable and complete, and a subset $\mathcal{A} \subset \mathcal{E}$ is pre-compact for $\mathrm{d}_{\mathrm{loc}}$ if and only if for each $r \geq 0$, the subset $[\mathcal{A}]_r \subset \mathcal{E}_r$ is pre-compact for δ_r (just reproduce the proof of Theorem 2). Here is an example of situation in which this more general setting is useful:

Exercise 6 (Pitman's Chinese restaurant process). For each $n \geq 1$, consider the space of permutations $\mathfrak{S}_n$ endowed with the trivial distance δ_n. For $n \geq 2$ and $\sigma \in \mathfrak{S}_n$, we define the restriction $[\sigma]_{n-1}^n \in \mathfrak{S}_{n-1}$ of σ as the permutation obtained by representing σ as a product of cycles with disjoint supports and removing n from the cycle in which it is present. Formally, we have

$$\left([\sigma]_{n-1}^n\right)(k) = \begin{cases} \sigma(k) & \text{if } \sigma(k) \neq n, \\ \sigma(n) & \text{otherwise,} \end{cases} \qquad \text{for all } k \in \{1, \ldots, n-1\}.$$

Check that the assumptions listed above are satisfied. In particular, the space $\mathcal{E}$ of coherent sequences of permutations endowed with the local topology is Polish. Is it possible to give another description of $\mathcal{E}$ in this particular example? Show that if $\boldsymbol{\sigma_n} \in \mathfrak{S}_n$ is a uniform random permutation, then $[\boldsymbol{\sigma_n}]_r^n \in \mathfrak{S}_r$ is a uniform random permutation, for $r \leq n$. Can you construct a random variable $\boldsymbol{\sigma} \in \mathcal{E}$ such that for each $n \geq 1$, the restriction $[\boldsymbol{\sigma}]_n \in \mathfrak{S}_n$ of $\boldsymbol{\sigma}$ is uniform? *Hint:* Pitman's Chinese restaurant process (see Wikipedia or [35, Section 10.1]).

1.2.2 Convergences of random variables with values in Polish spaces

Let us recall basics of convergence of random variables with values in a Polish space and then specialize it to the case of local topologies. We refer, e.g., to [22] for details and proof.

Let (E, d) be a Polish space. A random variable X with values in E is a measurable function from an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in (E, d) endowed with the Borel σ-field $\mathcal{B}$. If we have a sequence of random variables $(X_n)_{n\geq 1}$ and X defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ recall that

$$X_n \xrightarrow[n\to\infty]{(\mathbb{P})} X \iff \forall \varepsilon > 0, \quad \mathbb{P}(d(X_n, X) > \varepsilon) \xrightarrow[n\to\infty]{} 0$$

$$X_n \xrightarrow[n\to\infty]{(a.s.)} X \iff \mathbb{P}(X_n \xrightarrow[n\to\infty]{} X) = 1.$$

Those convergences, respectively called convergence in probability and almost sure convergences are the most common notion of convergence for coupled random variables (they need to be defined on the same probability space). Recall that a sequence of random variables with values in a Polish space converges in probability if and only if from any subsequence we can extract a further subsequence converging almost surely. But even if those random variables are defined on different probability spaces, we can speak of convergence of their laws:

$$X_n \xrightarrow[n\to\infty]{(d)} X \iff \mathcal{L}(X) \xrightarrow[n\to\infty]{weak} \mathcal{L}(X_n) \iff \forall f \in \mathcal{C}_b(E,\mathbb{R}), \quad \mathbb{E}[f(X_n)] \xrightarrow[n\to\infty]{} \mathbb{E}[f(X)],$$

where $\mathcal{L}(X)$ denotes the law of the random variable X, i.e. the image measure on E of $\mathbb{P}$ by the function X. This notion of convergence, called convergence in law or in distribution, is not a convergence of random variables, but rather a convergence of measures. This is arguably the most important notion of convergence in probability theory and it is implied by the convergence in probability which itself is implied by the almost sure convergence. The above three notions of convergence are stable under taking the image by a continuous function.

Recall that the set $\mathcal{M}_1(E)$ of all Borel probability measure on $(E, d, \mathcal{B})$ with the topology of weak convergence is itself a Polish space (it can be metrized e.g. by the Prokhorov distance). The standard pre-compactess criterion in Polish space in this context yields the notion of tightness of (laws of) random variables: A family $(X_i)_{i\in I}$ of random variables is tight if for any $\varepsilon > 0$ there exists a compact $A_\varepsilon \subset E$ such that for any $i \in I$ we have

$$\mathbb{P}(X_i \in A_\varepsilon) \geq 1 - \varepsilon.$$

We finally recall the Skorokhod embedding theorem which enables to transform a convergence in law into an almost sure convergence: Let $(\mu_n)_{n\geq 1}$ be a sequence of laws on $\mathcal{M}_1(E)$ so that $\mu_n \to \mu$ in distribution, then one can *construct* a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and random variables $X_n,\ X : \Omega \to E$ such that $\mathcal{L}(X) = \mu$ and $\mathcal{L}(X_n) = \mu_n$ for all $n \geq 1$ and such that

$$X_n \xrightarrow[n\to\infty]{a.s.} X.$$

Let us now focus on the case when the underlying Polish space (E, d) is equal to $(\mathcal{E}, \mathrm{d_{loc}})$. In our case, by the construction of the local topology it is easy to check:

Proposition 3. *Let $(X_i : i \geq 1)$ be a family of random variables with values in $(\mathcal{E}, \mathrm{d_{loc}})$.*

- $X_i \xrightarrow[i\to\infty]{} X$ *almost surely, in probability or in distribution, if and only if for any $r \geq 1$ we have $[X_i]_r \xrightarrow[i\to\infty]{} [X]_r$ almost surely, in probability or in distribution respectively.*

- *The family* $(X_i : i \geq 1)$ *is tight if and only if for any* $r \geq 1$ *the family* $([X_i]_r : i \geq 1)$ *is tight.*

Proof. For the first point, if $X_i \to X$ in probability or in distribution, since the restriction map $[\cdot]_r : (\mathcal{E}, \mathrm{d_{loc}}) \longrightarrow ([\mathcal{E}]_r, \delta)$ is continuous, we deduce that $[X_i]_r \to [X]_r$ in the same sense by the continuous mapping theorem. Conversely, suppose that for each $r \geq 0$, we have $[X_i]_r \to [X]_r$ in law as $i \to \infty$, and let us prove that $X_i \to X$ in law as $i \to \infty$. By [22, Theorem 2.1], it suffices to show that $\mathbb{E}[f(X_i)] \to \mathbb{E}[f(X)]$ as $i \to \infty$ for each *uniformly* continuous and bounded function $f : \mathcal{E} \to \mathbb{R}$. Noticing then that $\mathrm{d_{loc}}(x, [x]_r) \leq 2^{-r}$, we have

$$\begin{aligned} |\mathbb{E}[f(X_i)] - \mathbb{E}[f(X)]| \leq\ & \mathbb{E}[|f(X_i) - f([X_i]_r)|] \\ & + |\mathbb{E}[f([X_i]_r)] - \mathbb{E}[f([X]_r)]| \\ & + \mathbb{E}[|f([X]_r) - f(X)|], \end{aligned}$$

where the first and third term are bounded by the modulus of continuity of f at 2^{-r}, and the middle term tends to 0 as $i \to \infty$ by hypothesis. This proves the desired convergence in law. The convergence in probability is dealt with similarly.

For the second point, the direct implication is clear by continuous mapping. For the converse claim, for any $\varepsilon > 0$ one can find compact subsets $K_r \subset [\mathcal{E}]_r$ such that $\mathbb{P}([X_i]_r \in K_r) \geq 1 - \varepsilon 2^{-r}$ for all $i \geq 1$. We then set

$$K = \{x \in \mathcal{E} : [x]_r \in K_r \text{ for all } r \geq 1\} = \bigcap_{r \geq 1} [\cdot]_r^{-1}(K_r).$$

The sets $[\cdot]_r^{-1}(K_r)$ are closed by continuity of $[\cdot]_r$, so their intersection is closed as well. We can then check using our characterization in Theorem 2 of the pre-compact sets in $(\mathcal{E}, \mathrm{d_{loc}})$ that K is compact and $\mathbb{P}(X_i \in K) \geq 1 - \varepsilon$ as desired. □

Finally, let us present a down-to-earth condition for convergence in distribution with respect to the local topology, in the special case where $(\mathcal{E}, \delta)$ is discrete. Let us denote by $\mathcal{C}$ the collection of sets of the form

$$\{x \in \mathcal{E} : [x]_r = a_r\},$$

for $r \geq 0$ and $a_r \in [\mathcal{E}]_r$. Notice that $\mathcal{C}$ is a π-system (i.e, closed under finite intersections) that is contained in the Borel σ-algebra of $(\mathcal{E}, \mathrm{d_{loc}})$ (by the continuity of the restriction maps $[\cdot]_r : (\mathcal{E}, \mathrm{d_{loc}}) \to ([\mathcal{E}]_r, \delta)$). Moreover, we claim that any open subset O of $(\mathcal{E}, \mathrm{d_{loc}})$ can be written as a countable union of elements of $\mathcal{C}$. Indeed, contemplate the following identity:

$$O = \bigcup_{\substack{r \geq 0 \text{ and } a_r \in [\mathcal{E}]_r \\ \{x \in \mathcal{E} : [x]_r = a_r\} \subset O}} \{x \in \mathcal{E} : [x]_r = a_r\}. \tag{1.4}$$

The converse inclusion is automatic, now let us check the direct inclusion. Fix $x_0 \in O$, and fix $\varepsilon > 0$ such that $B_{\mathrm{d_{loc}}}(x_0, \varepsilon) \subset O$. Then, fix $R \geq 0$ large enough so that $2^{-R} < \varepsilon$, and let $a_R = [x_0]_R$. By construction, we have $x_0 \in \{x \in \mathcal{E} : [x]_R = a_R\}$. Moreover, this set is contained in $B_{\mathrm{d_{loc}}}(x_0, \varepsilon) \subset O$: for every $x \in \mathcal{E}$ such that $[x]_R = a_R$, we have $[x]_r = [x_0]_r$ for all $0 \leq r \leq R$, and thus $\mathrm{d_{loc}}(x, x_0) \leq \sum_{r \geq R+1} 2^{-r} = 2^{-R} < \varepsilon$. This proves (1.4), which concludes this argument. Back to our discussion, we deduce the following:

Proposition 4 (Local convergence in law, discrete case)**.** *A sequence of random variables $(X_n)_{n\geq 0}$ converges in distribution towards X_∞ if and only if for every $r \geq 0$ and every $a_r \in [\mathcal{E}]_r$ we have $\mathbb{P}([X_n]_r = a_r) \to \mathbb{P}([X_\infty]_r = a_r)$ as $n \to \infty$.*

Proof. The converse implication (which is the more useful) is granted by [22, Theorem 2.2], given the above discussion. For the direct implication, observe that the elements of $\mathcal{C}$ are both open and closed for $\mathrm{d_{loc}}$ (they are clopen): the result follows from [22, Theorem 2.1]. □

Beware, we may have the convergence of all $\mathbb{P}([X_n]_r = a_r)$ for $a_r \in [\mathcal{E}]_r$ without having the convergence in law of the (X_n)'s: indeed those probabilities may e.g. all tend to 0.

1.2.3 Two examples: random pointed graphs and plane trees

Let us now describe the discrete set of graphs and the restriction operations in more detail in two examples that will occupy us throughout the course.

Plane trees. We recall the formalism for plane trees as found in [58]. Let

$$\mathbb{U} = \bigcup_{n=0}^{\infty} (\mathbb{Z}_{>0})^n$$

where $\mathbb{Z}_{>0} = \{1, 2, \ldots\}$ and $(\mathbb{Z}_{>0})^0 = \{\varnothing\}$ by convention. An element u of $\mathbb{U}$ is thus a finite sequence of positive integers. We let $|u|$ be the length of the word u. If $u, v \in \mathbb{U}$, uv denotes the concatenation of u and v. If v is of the form uj with $j \in \mathbb{Z}_{>0}$, we say that u is the **parent** of v or that v is a **child** of u. More generally, if v is of the form uw, for $u, w \in \mathbb{U}$, we say that u is an **ancestor** of v or that v is a **descendant** of u.

Definition 3. *A (locally finite)* ***plane tree*** *τ is a (finite or infinite) subset of $\mathbb{U}$ such that*

1. *$\varnothing \in \tau$ ($\varnothing$ is called the* ***root*** *of τ),*
2. *if $v \in \tau$ and $v \neq \varnothing$, the parent of v belongs to τ*
3. *for every $u \in \mathbb{U}$ there exists $k_u(\tau) \in \{0, 1, 2, \ldots\}$ such that $uj \in \tau$ if and only if $j \leq k_u(\tau)$.*

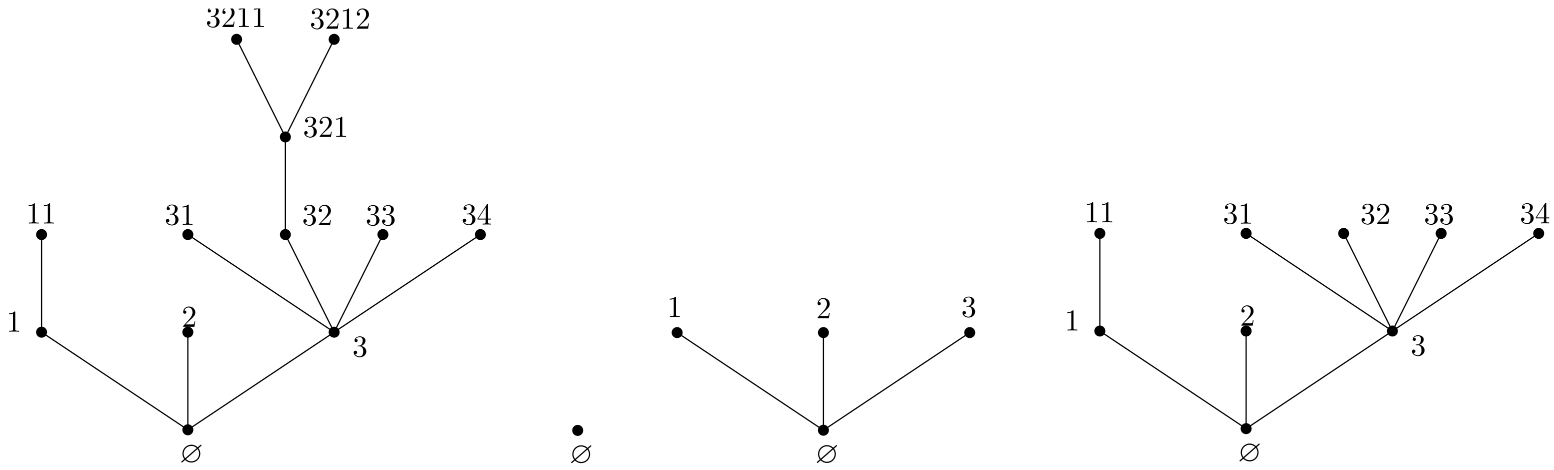


Figure 1.4: A finite plane tree and its restrictions of radius $0, 1, 2$.

A plane tree can be seen as a graph in which an edge links two vertices u, v whenever u is the parent of v or vice-versa. Notice that with our definition, vertices of infinite degree are not allowed since k_u cannot be infinite. This graph is of course a tree in the graph-theoretic sense, and has a natural embedding in the plane, in which the edges from a vertex u to its children $u1, \dots, uk_u(\tau)$ are drawn from left to right. All the trees considered in these pages are plane trees. The integer $|\tau|$ denotes the number of edges of τ and is called the size of τ.

We denote by $\mathcal{T}$ the set of all (finite or infinite) plane trees (sometimes called rooted plane trees). If $\tau \in \mathcal{T}$ is a plane tree, the notion of restriction $[\tau]_r$ of radius r we use is just the plane tree obtained by keeping all the vertices of τ which are at generation less than r from the origin, that is $[\tau]_r = \{u \in \tau : |u| \le r\}$. It is easy then to check that the conditions $(*)$ are in force, so that Theorem 2 gives rise to the local topology on $\mathcal{T}$, which is then Polish. A random plane tree will thus be seen as a random variable with values in $(\mathcal{T}, \mathrm{d}_{\mathrm{loc}})$.

Random pointed graphs.

Definition 4. *A pointed graph* $\mathsf{g}^\bullet$ *is a pair* (g, ρ) *where* g *is a (countable, locally finite, connected) graph and* $\rho \in \mathrm{V}(\mathsf{g})$ *is a reference vertex sometimes called the origin of the graph. Two pointed graphs* $\mathsf{g}^\bullet = (\mathsf{g}, \rho)$ *and* $\mathsf{h}^\bullet = (\mathsf{h}, \varrho)$ *are equivalent if there exists a graph homomorphism between* $\mathsf{g}^\bullet$ *and* $\mathsf{h}^\bullet$ *which sends* ρ *onto* ϱ *(we speak of pointed graph homomorphism).*

In what follows we will obviously identify equivalent graphs and so formally work on the space of equivalence classes of pointed graphs. We will implicitly make this identification and later speak of pointed graphs (instead of equivalence classes of pointed graphs). We introduce the set $\mathcal{G}^\bullet$ of all (equivalence classes) of (locally finite, countable, connected) pointed graphs. If $\mathsf{g}^\bullet$ is a pointed graph, we denote by $[\mathsf{g}^\bullet]_r$ the restriction of radius r

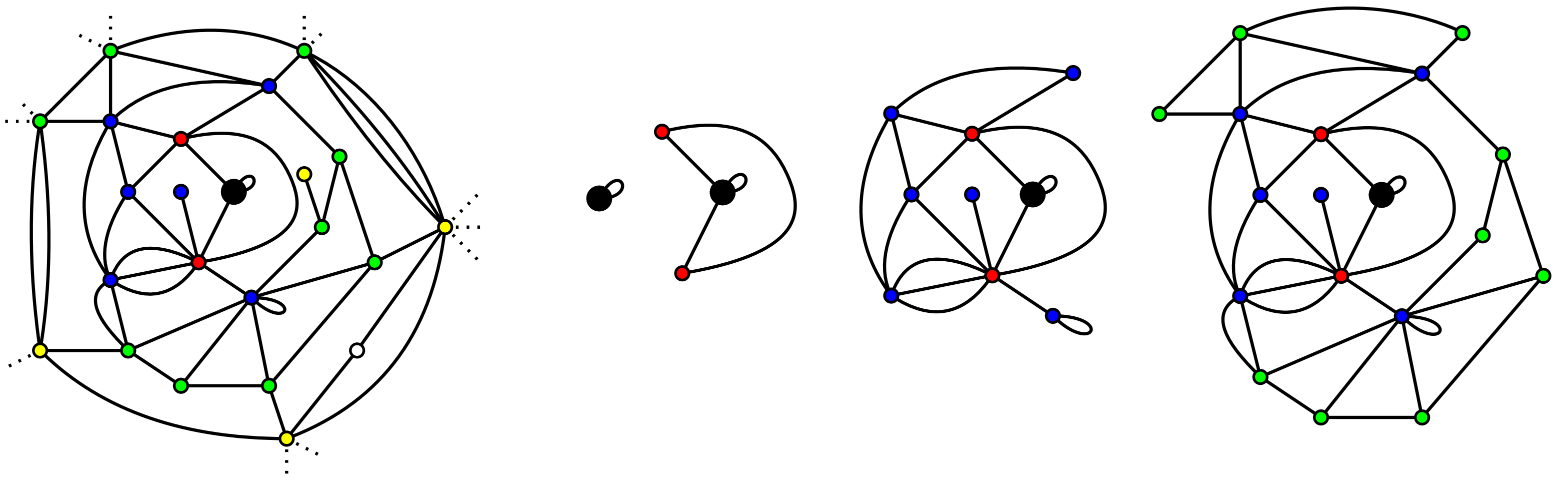


Figure 1.5: Illustration of the balls of radius $0, 1, 2, 3$ in a pointed graph.

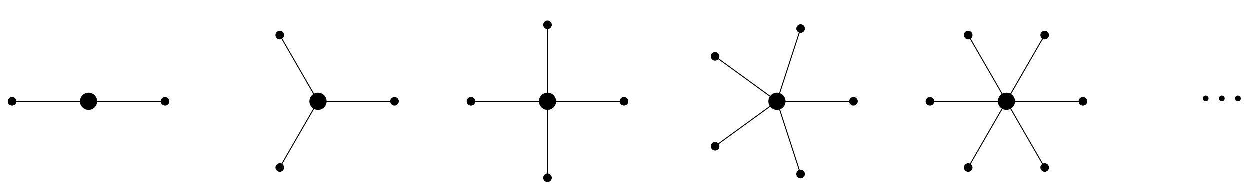
around the origin of $\mathbf{g}^\bullet$ to be the (equivalence class of the) graph obtained from $\mathbf{g}^\bullet = (\mathbf{g}, \rho)$ by keeping only those vertices which are at distance less than r from ρ and the edges between them; the resulting graph being pointed at ρ. See Figure 1.5

The compatibility relations of the restrictions are easy to check. The fact that there are only countably many restrictions of radius r is also easy to see since we restricted to locally finite graphs. It requires a bit of thought to show that if $\mathbf{g}_1^\bullet, \mathbf{g}_2^\bullet, \ldots$ is a sequence of compatible graphs in the sense that $[\mathbf{g}_j^\bullet]_r = \mathbf{g}_r^\bullet$ for $r \leq j$ then there exists a unique (equivalence class of a) infinite pointed graph $\mathbf{g}^\bullet$ whose restrictions of radius r is $\mathbf{g}_r^\bullet$ but this is also true. Hence we can apply Theorem 2 to endow $\mathcal{G}^\bullet$ with a local distance making it a Polish space. In this case, it is also easy to check that the pre-compact subsets $\mathcal{A} \subset \mathcal{G}^\bullet$ are those satisfying, for every $r \geq 0$:

$$\sup_{\mathbf{g}^\bullet \in \mathcal{A}} \max_{x \in \mathrm{V}([\mathbf{g}^\bullet]_r)} \deg(x) < \infty.$$

It follows that for any $M > 0$, the subset of pointed graphs in which the degree of any vertex is bounded by M is a compact set.

Exercise 7. Show that the following family of pointed graphs is not compact for the local topology on pointed graphs



Exercise 8. Show that a family of random pointed graphs $(G_i^\bullet)_{i \in I}$ is tight if and only if for any $r \geq 0$ the family of random variables

$$\max_{x \in \mathrm{V}([G_i^\bullet]_r)} \deg(x), \qquad i \in I$$

is tight as a family of real-valued random variables.

If g is a transitive graph, we can unambigously speak of its pointed version $\mathsf{g}^\bullet$ without specifying the pointed vertex since (g, x) is isomorphic to (g, y) for any $x, y \in \mathrm{V}(\mathsf{g})$. A random pointed graph will be, in these notes, a random variable $G^\bullet$ taking values in the Polish space $(\mathcal{G}^\bullet, \mathrm{d}_{\mathrm{loc}})$ endowed with the Borel σ-field.

Exercise 9 (Monster). Construct an infinite countable and locally finite graph M so that for any pointed graph $\mathsf{g}^\bullet \in \mathcal{G}^\bullet$, there exists a sequence of vertices $x_n \in \mathrm{V}(\mathsf{M})$ so that

$$(\mathsf{M}, x_n) \xrightarrow{n\to\infty} \mathsf{g}^\bullet.$$

In a sense, M "contains" all pointed locally finite graphs.

Bibliographical notes. Although it had many precursors, the formalization of the local topology for pointed graphs is due to Benjamini & Schramm in their pioneer paper [20]. The general framework presented here enables to treat local topologies on various spaces at once. It is also present in [66]. The formalism for plane trees has been introduced by Neveu [58], see [35, Chapter 4] for details.