Pure and Applied Functional Analysis
Volume 8, Number 6, 2023, 1767–1779

Yokohama Publishers
ISSN 2189-3764 ONLINE JOURNAL

TANGENT GRAPHS

PETER MÖRTERS

Dedicated to Prof David Preiss on the occasion of his 75th birthday

Abstract. The purpose of this note is to demonstrate an analogy of the concepts of weak local limits in the theory of random graphs and of tangent measures in geometric measure theory. We sketch two examples of random graphs for which weak local limits, or 'tangent graphs', can be used to explore their global regularity properties.

1. Introduction

1.1. **Motivation and background.** A key problem in geometric measure theory is to characterise global regularity of objects by means of local quantities. A great landmark in this endeavour is Preiss' Theorem [17]. In one version of this result the objects are Borel sets $E \subset \mathbb{R}^d$ of positive and finite k-dimensional Hausdorff measure

$$0 < \mathscr{H}^k(E) < \infty.$$

Regularity here means *partial k-rectifiability*, i.e. the existence of a Lipschitz map $f\colon \mathbb{R}^k \to \mathbb{R}^d$ such that

$$\mathscr{H}^k\big(E \cap f(\mathbb{R}^k)\big) > 0,$$

and the (necessary and sufficient) local criterion for regularity is the existence of a positive density on a set of positive measure, i.e.

$$\mathscr{H}^k\left(\left\{x \in E\colon 0 < \lim_{r\downarrow 0} \frac{\mathscr{H}^k(E \cap B(x,r))}{r^k} < \infty\right\}\right) > 0.$$

To prove this result Preiss introduced the concept of *tangent measures* as a key tool. The idea of a tangent measure at a point $x \in E$ is to zoom into the neighbourhood of x by looking at the measures $\mu_{x,r}$ defined by

$$\mu_{x,r}(A) = r^{-k}\mathscr{H}^k\big((x + rA) \cap E\big)$$

for small $r > 0$. The limit points of $(\mu_{x,r})$ in the vague topology of measures, as $r \downarrow 0$, are the tangent measures at x. The proof is based on characterising partial rectifiability in terms of properties that hold simultaneously for the tangent measures at all points in a set of positive measure and verifying that these properties are implied by the local criterion.

2020 *Mathematics Subject Classification.* 05C82, 28A80.

Key words and phrases. Tangent measure, rectifiability, random graph, preferential attachment, weak local limit, giant component, tangent.

More precisely, if the k-density exists and is positive and finite for all $x \in E'$ with $\mathscr{H}^k(E') > 0$, then for $\mathscr{H}^k$-almost every $x \in E'$ the set of tangent measures at x is nonempty and contains only k-uniform measures with $0 \in \operatorname{supp} \mu$, see [14, Proposition 3.4]. Here a measure μ is called *k-uniform* if, for some $c > 0$,

$$\mu(B(y,r)) = cr^k \text{ for every } y \in \operatorname{supp} \mu \text{ and } r > 0.$$

Additionally, at $\mathscr{H}^k$-almost every $x \in E'$ there exists a tangent measure which is *flat*, i.e. a constant multiple of $\mathscr{H}^k$ restricted to a k-dimensional subspace of $\mathbb{R}^d$, [14, Theorem 6.8]. These two properties of tangent measures imply partial k-rectifiability, see [14, Theorem 6.10].

Naturally, the situation in random graph theory is not as sophisticated and technically involved as in geometric measure theory. But still, the general concept of tangents is used to prove (or disprove) regularity in a somewhat analogous way: A property that holds for the tangents around all vertices taken from a set of positive measure implies a global regularity property. The role of tangents is now played by weak local limits of graphs, the 'tangent graphs', which were introduced by Benjamini and Schramm [3]. We now explain this analogy in detail.

The objects of interest are sequences of graphs $(G^{(N)})_N$ such that $G^{(N)}$ has N vertices. We study the limit $N \to \infty$ for sparse sequences, which means that the number of edges is also of order N. The graphs may or may not be random, and they need not be embedded into each other or into Euclidean space. The regularity that we are interested in is the existence of a *giant component*, i.e. a connected subgraph $C^{(N)} \subset G^{(N)}$ such that

$$\liminf_{N\to\infty} \frac{|C^{(N)}|}{N} > 0.$$

This means that there exist $\theta > 0$ such that (with probability converging to one) a proportion θ of the graph can be covered by a connected subgraph, so that connectedness of a graph plays a similar role here as being the image of a Lipschitz function in analysis.

There are two principal differences between the measure theoretical and the graph theoretical set-up that are crucial when defining an analogue to the notion of tangent measures. *First*, as the minimal metric distances in graphs are one, rather than 'zooming into a point' we need to keep the observation scale fixed (or even let it grow slowly) while the surrounding graph grows. *Second*, and even more important, there is no natural global limiting object of the graphs from which we could choose the point at which we take the tangent. Instead, we pick a random point from each graph $G^{(N)}$. Our tangents therefore do not represent what a concrete object looks like around a specific point but rather what the statistics of objects is around a typical point in a large graph.

Looking at the regularity results in geometric measure theory the local criteria implicitly refer to the neighbourhood of a 'randomly chosen' point from the set E. We can therefore uphold the analogy when defining a tangent based on neighbourhoods of randomly chosen vertices in $G^{(N)}$. Limits now have to be taken on the probability space of neighbourhoods, i.e. in the weak or distributional sense. The notion of a tangent graph is therefore statistical (even if the graphs are not random) in the sense that it is a random object that tells us what a neighbourhood of a typical point looks like on average as the graphs get large. This often suffices to get information about the (lack of) regularity of the graphs.

1.2. **Weak local limits.** We now carefully introduce our concept of tangent for graph sequences $(G^{(N)})_N$. The tangents will be probability measures on the space $\mathscr{G}_*$ of locally finite, rooted graphs. Elements (G, o) of $\mathscr{G}_*$ are graphs G with a finite or countably infinite vertex set and a distinguished vertex $o \in G$ called the root, such that every vertex has a finite number of adjacent edges (called its degree). For $(G, o) \in \mathscr{G}_*$ and $n \in \mathbb{N}$ we denote by $G \wedge n$ the finite subgraph consisting of all vertices in G that can be reached from o by a path with no more than n edges. Two elements $(G_1, o_1), (G_2, o_2) \in \mathscr{G}_*$ have metric distance $d(G_1, G_2) = \frac{1}{N+1}$ where

$$N = \max \big\{ n \colon \exists \text{ bijection } \phi \colon G_1 \wedge n \to G_2 \wedge n \text{ with } \phi(o_1) = o_2 \text{ and}$$

$$\{u, v\} \text{ is an edge in } G_1 \text{ iff } \{\phi(u), \phi(v)\} \text{ is an edge in } G_2 \big\}.$$

We identify (G_1, o_1) and (G_2, o_2) if $N = \infty$. In particular, we identify a rooted graph with the connected component of its root. With this convention $\mathscr{G}_*$ becomes a complete, separable metric space.

We take a sequence of graphs $(G^{(N)})_N$ such that $G^{(N)}$ has N vertices and turn $G^{(N)}$ into a random rooted graph by picking the root uniformly at random from the vertices of $G^{(N)}$. We say that $(G^{(N)})_N$ *converges weakly locally* to a random rooted graph (G, o) if, for every bounded, continuous function $h \colon \mathscr{G}_* \to \mathbb{R}$ we have

$$\frac{1}{N} \sum_{o \in G^{(N)}} h\big(G^{(N)}, o\big) \overset{N \to \infty}{\longrightarrow} \mathbb{E}\big[h(G, o)\big].$$

If the graphs $(G^{(N)})_N$ themselves are random, the objects on the left hand side are random variables and the convergence is supposed to hold in probability. A useful tool to establish weak local convergence is that it has to be checked only for functions h which are the indicator function of balls in $\mathscr{G}_*$.

An example of an interesting function h to which weak local limits apply is given by $h((G, o)) = 1/|C_o|$, where $|C_o|$ is the number of vertices in the connected component of o if this finite, and $h((G, o)) = 0$ otherwise. Then the l.h.s. is the number of connected components in the finite graph $G^{(N)}$ divided by N and weak local convergence implies that this random quantity converges in probability to a constant which is positive if and only if the limiting rooted graph is finite with positive probability.

The weak local limit (G, o), if it exists, plays the role of a tangent graph. In the remainder of this section we compare characteristic properties of tangent measures and of tangent graphs. In the main part of this paper we then identify the tangent graph of two random graph models and demonstrate its use in identifying the regularity of the underlying graph sequence.

1.3. **Regularity of tangent measures.** Key to the usefulness of any concept of tangent is that the tangents to an object should have better symmetry or invariance properties than the original object. One way to make the invariance properties of tangent measures visible is to randomise the scales at which we zoom into the neighbourhoods of $x \in E$. To this end we look at the random measure $\mu_{x,R}$ where R is chosen randomly according to the truncated Haar measure

$$\frac{\mathbb{1}_{\delta<r<1}}{\log(1/\delta)}\frac{dr}{r}.$$

The distributional limit points of $\mu_{x,R}$ as $\delta \downarrow 0$ are random measures with values in the set of tangent measures at x, they (or, more precisely, their distributions) are called *tangent measure distributions*, see [9,15]. Loosely speaking, tangent measure distributions describe what we see at *random* scales when we zoom into x.

The invariance of tangent measure distributions can be expressed using the concept of *Palm distributions*. Loosely speaking, these are derived from stationary distributions of measures by conditioning on the origin being in the support of the measure. More specifically, a σ-finite measure $\mathbb{Q}$ on the space of measures is stationary if $\mathbb{Q} \circ \theta_x^{-1} = \mathbb{Q}$ for all $x \in \mathbb{R}^d$ where the shift θ_x is given by $\theta_x\mu(B) = \mu(x+B)$. For a measurable function $w\colon \mathbb{R}^d \to [0,\infty)$ with $\int w(x)\,dx = 1$ we define the intensity of $\mathbb{Q}$ as

$$\lambda_{\mathbb{Q}} = \iint w(x)\,\mu(dx)\,\mathbb{Q}(d\mu),$$

which does not depend on the choice of w. If the intensity is finite, the Palm distribution $\mathbb{P}$ associated with $\mathbb{Q}$ is given by

$$\mathbb{P}(A) = \lambda_{\mathbb{Q}}^{-1} \iint \mathbb{1}_{\theta_x\mu\in A} w(x)\,\mu(dx)\,\mathbb{Q}(d\mu).$$

A probability measure $\mathbb{P}$ is called a *Palm distribution* if it is the Palm distribution associated with some stationary measure of finite intensity. Palm distributions are characterised by the invariance equation

$$\iint G(\mu,u)\,\mu(du)\,\mathbb{P}(d\mu) = \iint G(\theta_u\mu,-u)\,\mu(du)\,\mathbb{P}(d\mu), \tag{1.1}$$

which holds for all $G\colon \mathscr{M}(\mathbb{R}^d)\times\mathbb{R}^d \to [0,\infty)$, see [19]. It is shown in [15] that at $\mathscr{H}^k$-almost every $x \in E$ all tangent measure distributions are Palm distributions.

1.4. **Regularity of weak local limits.** Equation (1.1) can be read as a mass transport principle for tangent measures. A similar property holds for weak local limits, it has been introduced by Benjamini and Schramm [3], see also [1] for an extensive survey.

We take $\mathscr{G}_{**}$ to be the space of graphs with two distinguished vertices in the same connected component. Up to isomorphism classes this can be turned into a metric space in a similar way as $\mathscr{G}_*$. A random rooted graph (G,o) is *unimodular* if it satisfies the mass transport principle, i.e.

$$\int \sum_{x\in G} g(G,o,x)\mathbb{P}(d(G,o)) = \int \sum_{x\in G} g(G,x,o)\mathbb{P}(d(G,o)), \tag{1.2}$$

for every Borel measurable $g\colon \mathscr{G}_{**} \to [0,\infty]$ where the sums are over all x in an arbitrary connected graph G representing the class (G,o). To understand the role of the unimodularity put $g(G,x,y) = f(\deg(x))/|G|$ where $\deg(x)$ is the degree of x in G. If the graph is finite almost surely the l.h.s. in (1.2) is the expectation of $f(\deg(o))$ while the r.h.s. is the expectation of $f(\deg(x))$ for x uniformly chosen in G. This characterises the root as a typical vertex.

Benjamini and Schramm [3] have shown that any weak local limit is unimodular. The converse problem, whether any unimodular random rooted graph is a weak local limit, is an open problem. Aldous and Lyons [1] explain potential consequences of an affirmative answer.

2. Simple preferential attachment graphs

In this section we discuss a random graph model which is designed as possibly the simplest model sufficiently nontrivial to test the theory on it. The *simple preferential attachment graph* is an iteratively constructed sequence of random graphs where in each step a new vertex is added to the graph and connected to existing vertices independently with a preference for older (and implicitly more powerful) vertices. The idea of constructing graphs from this (or a similar) principle goes back to [2], see [12] for a more mathematical overview of different models. The arguments presented here are boiled down from the paper [6] in which we discuss a 'serious' preferential attachment model. The proofs in [6] are technically much harder but the ideas are the same in principle.

2.1. **The model.** Vertices arrive one-by-one and vertex n attaches to each earlier vertex $m \in \{1,\dots,n-1\}$ independently with a probability proportional to $m^{-\gamma}$ for some parameter $0<\gamma<1$ parametrising the strength of the preference of early vertices. The proportionality factor is chosen so that the expected number of connections of a vertex is asymptotically constant. As, for some constant $c>0$,

$$\sum_{m=1}^{n-1} m^{-\gamma} \sim cn^{1-\gamma}$$

the proportionality factor has to be of order $n^{\gamma-1}$. Altogether, the connection probability of two distinct vertices with number (or rank) i and j is

$$p_{ij} = \beta(i \vee j)^{\gamma-1}(i \wedge j)^{-\gamma}$$

for some fixed parameter $\beta>0$, and all connections are independent. The resulting random graph on the vertices $\{1,\dots,N\}$ is denoted $G^{(N)}$.

2.2. **The tangent graph.** The problem at hand is to find for which parameters $\beta>0$ and $0<\gamma<1$ there exists a giant component of $(G^{(N)})_N$. This problem can be explicitly solved, the main tool in our solution is to identify the weak local limit of this random graph sequence. To prepare this we now review an important concept from probability theory.

Given a σ-finite measure space $(X,\mathcal{X},\mu)$ we can always define a *Poisson process P with intensity measure μ*, which is a random measure on $(X,\mathcal{X})$ such that

- for every $A \in \mathcal{X}$ with $\mu(A) < \infty$ the random variable $P(A)$ is Poisson-distributed with expectation $\mu(A)$,
- for $A, B \in \mathcal{X}$ disjoint, the random variables $P(A)$ and $P(B)$ are independent.

Note that the random measure P takes values in the nonnegative integers almost surely, i.e. it is purely atomic. The idea is that the Poisson process randomly scatters points on X and $P(A)$ counts the points landing in A. It is easy to see that for every measurable $f: X \to [0, \infty)$ we have

$$\mathbb{E}\Big[\int f\,dP\Big] = \int f\,d\mu.$$

We now construct a random rooted tree using Poisson processes to determine offspring numbers. For its construction vertices are given a position on the negative halfline $(-\infty, 0)$. We start with a root placed at a fixed position $-x \in (-\infty, 0)$. We sample a Poisson process P with intensity measure μ and place the children of the root such that for every Borel set $A \subset (-\infty, 0)$ the number of children in A is $P(x + A)$. In other words, when P has atoms $y_1, y_2, \ldots$ the first generation in the tree consists of vertices in positions $y_1 - x, y_2 - x, \ldots$ if these are negative (and otherwise they are discarded).

Continuing iteratively, given individuals in positions $-x_1, -x_2, \ldots, -x_k$ in the nth generation, we determine the next generation by sampling independent Poisson processes $P_1, P_2, \ldots, P_k$ and determining that the number of children of the individual at $-x_i$ which are placed in $A \subset (-\infty, 0)$ is $P_i(x_i + A)$. The process thus constructed is the *Poisson branching random walk with intensity measure μ and start in $-x$ killed at the origin.* Note that the expected number of children of a vertex in position $-x$ is $\mu((-\infty, x))$ so that the further to the left a particle is placed on the line, the more powerful it is, i.e. the more offspring it is expected to produce.

Forgetting the positions of the vertices in the branching random walk after its construction we obtain a random rooted tree denoted by $\mathfrak{T}(x, \mu)$. If

$$\mu((-\infty, a]) < \infty \text{ for all } a > 0$$

all the vertices of this tree have finite degree and hence $\mathfrak{T}(x, \mu)$ is a random element of the space $\mathscr{G}_*$ of locally finite, rooted graphs. The random rooted graph $\mathfrak{T}(x, \mu)$ could be finite or infinite and we will later see a sharp criterion for when it is almost surely finite.

Lemma 2.1. *The weak local limit of the simple preferential attachment graphs is the tree $\mathfrak{T}(X, \pi)$ associated with the Poisson branching random walk killed at the origin with initial particle placed at $-X$, where X is standard exponentially distributed, and intensity measure π given by*

$$\pi(dz) = \beta(e^{z(1-\gamma)} 1_{z<0} + e^{z\gamma} 1_{z>0})\,dz.$$

We now explain the idea behind this lemma. The vertices of $G^{(N)}$ are indexed $1, 2, \ldots, N$ with small indices indicating powerful vertices. Define

$$t_n = \sum_{k=1}^{n-1} \frac{1}{k} \sim \log n$$

and the mapping

$$\phi_N \colon \{1, \ldots, N\} \to (-\infty, 0], \quad n \mapsto t_n - t_N.$$

A uniformly chosen vertex is then mapped into a position which, as $N \to \infty$, converges to the law of $\log U$ for U uniform on $(0,1)$, which equals $-X$ for X standard exponentially distributed.

We now fix a relative age $u \in (0,1)$ and look at a vertex $\lceil uN \rceil \in G^{(N)}$ and the point process consisting of all points $\phi_N(j)$ such that the edge $\{j, \lceil uN \rceil\}$ is present. The claim is that this process converges in distribution to $\log u + P$ defined by

$$\log u + P(A) := P(-\log u + A) \text{ for } A \subset (-\infty, 0),$$

where P is a Poisson point process with intensity measure π.

Take $a < b < u$, then

$$\begin{aligned}
&\mathbb{P}\big(\lceil uN \rceil \text{ does not connect to any } j \in [aN, bN]\big) \\
&\quad = \prod_{j=\lceil aN \rceil}^{\lfloor bN \rfloor} \big(1 - \beta (uN)^{\gamma-1} j^{-\gamma}\big) \sim \exp\Big(-\beta (uN)^{\gamma-1} \sum_{j=\lceil aN \rceil}^{\lfloor bN \rfloor} j^{-\gamma}\Big) \\
&\quad \longrightarrow \exp\Big(-\beta u^{\gamma-1} \int_a^b x^{-\gamma}\, dx\Big) = \exp\Big(-\beta \int_{\log a - \log u}^{\log b - \log u} e^{z(1-\gamma)}\, dz\Big) \\
&\quad = \exp\big(-\pi(\log a - \log u, \log b - \log u)\big).
\end{aligned}$$

Hence the probability that there are no points $\phi_N(j) \in [\log a, \log b]$ with $j \sim \lceil uN \rceil$ converges to the probability that $\log u + P([\log a, \log b]) = 0$ and, by the same calculation, the expected number of points $\phi_N(j) \in [\log a, \log b]$ with $j \sim \lceil uN \rceil$ converges to the expectation of $\log u + P([\log a, \log b])$ as well. By Kallenberg's theorem, see e.g. [18, Proposition 3.22], this implies convergence to the Poisson process $\log u + P$ on $(-\infty, \log u]$ and a similar calculation gives convergence on $(\log u, 0]$.

Suppose we have explored the offspring of $\lceil uN \rceil$ and move to the next generation. If $\lceil vN \rceil$ is an offspring vertex, then we can do the same calculation omitting the vertices that we have already seen in the exploration so far. If these are $o(N)$ many we also get offspring distributed like an independent point process with distribution $\log v + P$.

For the actual proof of weak local convergence we fix a rooted tree t of fixed depth at most k and let h be the indicator of a ball centred in t with radius $\frac{1}{k+1}$. Refining the calculation above we can construct $\mathfrak{T}(\log u, \pi)$ and the rooted graph $(G^{(N)}, \lceil uN \rceil)$ on the same probability space such that

$$h\big(G^{(N)}, \lceil uN \rceil\big) \to h\big(\mathfrak{T}(\log u, \pi)\big) \text{ in probability.}$$

Such a construction is called a *coupling*. To complete the proof we then verify that the local neighbourhoods in different points of $G^{(N)}$ are sufficiently independent so that by a law of large numbers the average over all possible roots converges to the expectation of the limit under a uniformly chosen $u \in (0,1)$. Recall that if U is a uniform random variable on $(0,1)$, then $\log U = -X$ for a standard exponentially distributed X.

Being a random rooted tree associated with a 'classical' branching process the 'tangent graph' thus constructed for the simple preferential attachment graph is a much more accessible object than the original graph sequence.

2.3. **Existence of a giant component.** We now argue how a solution to our problem can be obtained from the weak local limit. Loosely speaking, a giant component exists if the weak local limit has positive probability of being an infinite connected graph. Indeed, let h_k be the indicator of the event that there is no self-avoiding path of length k starting at the root and denote the weak-local limit (G, o). Then, denoting by $C_v^{(N)}$ the connected component containing the vertex v,

$$\frac{1}{N}\#\{v \in G^{(N)} : \mathrm{diam}(C_v^{(N)}) \leq 2(k-1)\} \geq \frac{1}{N}\sum_{o \in G^{(N)}} h_k\big(G^{(N)}, o\big)$$

$$\longrightarrow \mathbb{P}((G, o) \text{ has no self-avoiding path of length } k \text{ starting at } o),$$

and as $k \to \infty$ the right hand side converges to the probability that (G, o) is a finite rooted graph. In particular, if (G, o) is a finite rooted graph almost surely, then $(G^{(N)})$ has no giant component. For the converse direction, i.e. to show that there is a giant component if the tangent graph has positive probability of being infinite, the convergence type in the weak local limit is not strong enough. In the present example this can be resolved by taking suitable sequences $h_n \to h$ of functions on $\mathscr{G}_*$ that capture growing graph neighbourhoods and replace the coupling argument above by

$$h_N\big(G^{(N)}, \lceil uN \rceil\big) \to h\big(\mathfrak{T}(\log u, \pi)\big) \text{ in probbability,}$$

see [6], but there is also a general criterion of van der Hofstad [11] that can be checked. This is more technical and we cannot give further details here.

To answer our initial question we still need to decide whether there is a positive probability that the tangent graph is an infinite rooted graph. The killed branching random walk is simple enough to explicitly decide when this is the case in terms of the model parameters $\beta > 0, 0 \leq \gamma < 1$.

Theorem 2.2. *In the simple preferential attachment model a giant component exists if and only if*

$$\gamma \geq \frac{1}{2} \text{ or } \beta > \frac{1}{4} - \frac{\gamma}{2}.$$

We can use general branching process theory [4] to find the parameters for which the tangent graph is finite. Observe that unimodularity implies that, if the tree is almost surely finite, the degree of a randomly chosen vertex has the same distribution as the degree of the root, and hence its offspring number is stochastically smaller and its spatial position biased to the right compared to the root. This is reflected in the finiteness criterion.

Indeed, the Poisson branching random walk killed at the origin with intensity π dies out (and hence the associated tree is finite) almost surely if and only if there exists $\alpha > 0$ such that

$$\rho(\alpha) := \int_{-\infty}^{\infty} e^{-\alpha t} \pi(dt) \leq 1.$$

This can be calculated as

$$\rho(\alpha) = \beta\Big(\int_{-\infty}^{0} e^{-\alpha z + (1-\gamma) z}\, dz + \int_{0}^{\infty} e^{-\alpha z + \gamma z}\, dz \Big),$$

which is finite only if $\gamma < \alpha < 1 - \gamma$. Such a choice is only possible if $\gamma < \frac{1}{2}$ and in this case $\rho(\alpha)$ equals

$$\frac{\beta}{1 - \alpha - \gamma} + \frac{\beta}{\alpha - \gamma}.$$

This expression is minimal if $\alpha = \frac{1}{2}$ and then equals $\frac{4\beta}{1-2\gamma}$. Hence no giant component exists if and only if $\gamma < \frac{1}{2}$ and $\beta \leq \frac{1}{4} - \frac{\gamma}{2}$.

In particular if $\gamma \geq \frac{1}{2}$ then a giant component exists even if the edge density (i.e. the total number of edges per vertex, which is proportional to β in our model) is arbitrarily small. For this robustness phenomenon the threshold of $\gamma = \frac{1}{2}$ (the value at which the variance of the degree distribution becomes infinite) is typical for graph models without clustering, which is the case when the tangent graph is a tree. However, this turns out differently in the spatial models we consider below.

3. Spatial preferential attachment graphs

The fact that the scale-free networks we looked at so far are locally tree-like was useful for the analysis, because we have a good understanding of random trees and, in particular, branching processes. But precisely this feature is not desirable for modelling purposes, as network data typically has clustering, i.e. the probability that two vertices in the neighbourhood of a given vertex are connected is much larger than the probability that two arbitrary vertices are connected. Spatial models address this problem: Points are now embedded in space and the probability of edges connecting two points depends on their geometric distance, in that far away points are much less likely to be connected. This creates the desired clustering effect in a natural way, but makes the models harder to analyse.

3.1. **The model.** We now discuss a nontrivial model of a spatial random graph with preferential attachment. The model is a variant of the model introduced and studied in [7] and thoroughly investigated in [8]. It is a substantially simplified version of the spatial preferential attachment model studied in [10, 13], for which many open problems remain.

Let $\mathbb{T}_1 = (-1/2, 1/2]^d$ be the d-dimensional torus of side-length one, endowed with the torus metric d. Let $G^{(1)}$ be the graph with a single vertex placed uniformly at random in $\mathbb{T}_1$ and no edges. Given the graph $G^{(N-1)}$ with vertices in positions $x_1, \ldots, x_{N-1}$ vertex N is placed in position x_N chosen uniformly at random in $\mathbb{T}_1$

and connected by an edge to each existing node $n \in \{1, \dots, N-1\}$ independently with probability

$$\rho\big(p_{nN}^{-1}\, \mathrm{d}(x_n, x_N)^d\big),$$

where p_{nN} are exactly as in the previous model, i.e. for $n < N$,

$$p_{nN} = \beta N^{\gamma-1} n^{-\gamma},$$

and $\rho\colon [0,\infty) \to [0,1]$ is a nonincreasing *profile function* standardised to satisfy

$$\int_{\mathbb{R}^d} \rho(|x|^d)\, dx = 1.$$

With this choice of parametrisation the probability that the Nth vertex connects to n is asymptotically p_{nN} as $N \to \infty$. Hence the connection probabilities are similar to the simple preferential attachment model but of course the correlation structure is much more complicated.

We can use the profile function ρ to weaken these correlations, i.e. the slower the asymptotic decay of ρ the weaker the correlations. Another point of view is that with ρ we can tune the influence of the geometry independently from the degree structure of the graph. Lighter tails of ρ mean stronger geometric restrictions, which are most rigid if $\rho(x) = \mathbb{1}_{[0,r]}(x)$ for $r > 0$. In this case vertex N is linked to vertex n if their positions are within distance $(rp_{nN})^{1/d}$.

For our most interesting results we take $\rho(x) \sim cx^{-\delta}$ for some parameter $\delta > 1$ as in this case the behaviour changes qualitatively as δ varies. Other choices of ρ can be studied by the same method.

3.2. **The tangent graph.** We now describe what turns out to be the weak local limit of our spatial preferential attachment network. Despite the fact that this limit is not a tree, it is still a useful and fascinating object, the age-dependent random connection model, which we now define.

Let $\mathcal{X}$ denote the Poisson process on $\mathbb{R}^d \times (0,1)$ with Lebesgue intensity measure. We denote the points supporting the Poisson process by $\mathbf{x} = (x, s)$ and say that $\mathbf{x}$ is born at time s and placed at position x. Given $\mathcal{X}$ we let $\mathcal{X}_0 = \mathcal{X} \cup \{\mathbf{o}\}$ where $\mathbf{o} = (0, U)$ is a vertex at the origin and U is an independent uniform mark in $(0,1)$. We define the random rooted graph $(G, \mathbf{o})$ by placing an edge between $\mathbf{x} = (x, u)$ and $\mathbf{y} = (y, s)$ with $s < u$ independently of everything else with probability

$$\rho\big(\beta^{-1} s^{\gamma} u^{1-\gamma}\, |x - y|^d\big). \tag{3.1}$$

If $\gamma < 1$ the rooted graph $(G, \mathbf{o})$ is almost surely locally finite. We call this random rooted graph the *age-dependent random connection model* with profile function ρ and parameters γ, β.

Lemma 3.1. *The weak local limit of the spatial preferential attachment graph is the age-dependent random connection model with the same profile function ρ and parameters γ, β.*

We now sketch the proof of this lemma. We map the points of $(G^{(N)})$ with positions $x_1, \dots, x_N$ onto $\mathbb{R}^d \times (0,1]$ via the mapping

$$h_N \colon \mathbb{T}_1 \times \{1, \dots, N\} \to \mathbb{R}^d \times (0,1], (x,n) \mapsto (N^{1/d}x, \tfrac{n}{N}).$$

where we identify the torus with $[-\frac{1}{2}, \frac{1}{2}]^d$ so that a randomly chosen vertex of $G^{(N)}$ is positioned at the origin. The index i of this vertex is uniformly chosen from $\{1, \dots, N\}$ and hence this vertex is mapped by h_N to a point placed at the origin with a mark which, as $N \to \infty$, is asymptotically uniformly distributed. The point process formed by the image of the other vertices

$$\{h_N(x_1, 1), \dots, h_N(x_N, N)\} \setminus \{h_N(x_i, i)\}$$

converges to the Poisson process on $\mathbb{R}^d \times (0,1)$. To see this we look at the probability that the compact set $A \times [a,b] \subset \mathbb{R}^d \times (0,1)$ contains no points, which is

$$\prod_{\substack{j=aN \\ j \neq iN}}^{bN} \big(1 - \mathrm{Vol}(N^{-1/d}A)\big) \longrightarrow \exp\big(-(b-a)\mathrm{Vol}(A)\big).$$

The limit is the probability that the Poisson process has no point in $A \times [a,b] \subset \mathbb{R}^d \times (0,1)$. Similarly we see that the expected number of points in $A \times [a,b]$ converges to $(b-a)\mathrm{Vol}(A)$. This shows that the point processes converge as claimed. Moreover, there is an edge between (x_n, n) and (x_m, m) for $n < m$ with probability

$$\begin{aligned} \rho\big(p_{nm}^{-1}\, \mathrm{d}(x_n, x_m)^d\big) &= \rho\big(\beta^{-1} m^{1-\gamma} n^{\gamma}\, \mathrm{d}(x_n, x_m)^d\big) \\ &= \rho\big(\beta^{-1} \big(\tfrac{n}{N}\big)^{\gamma} \big(\tfrac{m}{N}\big)^{1-\gamma} |N^{1/d}x_n - N^{1/d}x_m|^d\big), \end{aligned}$$

if $x_n, x_m \in B(x_i, \frac{1}{4})$ so that the distance in the torus metric and in the euclidean metric of the unit cube coincide. This connection probability agrees with the probability that $h_N(x_n, n)$ and $h_N(x_m, m)$ are connected by an edge. As, for every k, the probability that x_i is connected in k steps to a vertex outside $B(x_i, \frac{1}{4})$ goes to zero, we again have couplings of the neighbourhood of a randomly chosen vertex in $G^{(N)}$ to the tangent graph and the proof of Lemma 3.1 can be completed by a suitably tailored law of large numbers, see [16] for a statement that can be adapted to our requirements.

The Poisson process on $\mathbb{R}^d \times (0,1)$ (with Lebesgue intensity) is stationary under spatial shifts and the associated Palm distribution is simply the Poisson process on $\mathbb{R}^d \times (0,1)$ with an additional uniformly marked point at the origin. Hence our tangent graph is embedded in euclidean space to have the same mass transport property as the tangent measure distributions.

3.3. **Existence of a giant component.** Recall that we identify two rooted graphs if the connected components of the root agree, so that infiniteness of a rooted graph really refers to infiniteness of the connected component of the root. As before the spatial preferential attachment graph has no giant component if the tangent graph is almost surely finite, but some extra work is needed to show that the spatial preferential attachment graph has a giant component if the tangent graph is infinite with positive probability. This is proved for a variant of the model in the PhD thesis [14].

Unfortunately, the tangent graph is not easy to analyse and we cannot expect to find the precise range of γ, β and the tail index $\delta > 1$ of the profile function, where the rooted graph is infinite with positive probability. But we are able to identify the values of γ and δ so that for all $\beta > 0$ the rooted graph is infinite with positive probability.

Theorem 3.2. *In the spatial preferential attachment model a giant component exists if*

$$\gamma > \tfrac{\delta}{\delta+1} \text{ and } \beta > 0 \text{ arbitrary;}$$

but if $\gamma < \frac{\delta}{\delta+1}$ no giant component exists for all sufficiently small $\beta > 0$.

This is proved in the PhD thesis [14] and in the paper [4]. The most difficult step is to check the infiniteness of the connected component of the origin in the age-dependent random connection model.

We note that this result is different from Theorem 2.2 because in that case for all $\gamma > \frac{1}{2}$ and $\beta > 0$ a giant component exists. The geometric embedding and correlation structure influences the graph topology by making it less connected. If $\delta > \frac{\gamma}{1-\gamma}$ the constraints coming from the geometry are strong enough so that a giant component can only emerge if β is sufficiently large.

4. Conclusion

We have seen that the idea of weak local limits used in the study of random graphs has a strong conceptual similarity with the idea of tangent measures used in geometric measure theory. It is my hope that this note can inspire a further transfer of ideas between these mathematical fields.

Acknowledgment

I would like to thank Christian Mönch for bringing reference [1] and particularly the general concept of unimodularity to my attention. I would also like to thank an anonymous referee for helpful comments.

References

[1] D. Aldous and R. Lyons, *Processes on unimodular random networks*, Electron. J. Probab. **12** (2007), 1454–1508.

[2] A.-L. Barabasi and R. Albert, *Emergence of scaling in random networks*, Science **286** (1999), 509–512.

[3] I. Benjamini and O. Schramm, *Recurrence of distributional limits of finite planar graphs*, Electron. J. Probab. **6** (2001), 1–13.

[4] J. D. Biggins, , B.D. Lubachevsky, A. Shwartz and A. Weiss, *A branching random walk with a barrier*, Ann. Appl. Probab. **1** (1991), 573–581.

[5] C. De Lellis, *Rectifiable Sets, Densities and Tangent Measures*, Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2008.

[6] S. Dereich and P. Mörters, *Random networks with sublinear preferential attachment: The giant component*, Ann. Probab. **41** (2013), 329–384.

[7] P. Gracar, A. Grauer, L. Lüchtrath and P. Mörters, *The age-dependent random connection model*, Queueing Syst. **93** (2019), 309–331.

[8] P. Gracar, L. Lüchtrath, P. Mörters, *Percolation phase transition in weight-dependent random connection models*, Adv. Appl. Prob. **53** (2021), 1090–1114.
[9] S. Graf, *On Bandt's tangential distribution for self-similar measures*, Monatshefte für Mathematik **120** (1995), 223–246.
[10] C. Hirsch and C. Mönch, *Distances and large deviations in the spatial preferential attachment model*, Bernoulli **26** (2020), 927–947.
[11] R. van der Hofstad, *The giant in random graphs is almost local*, Preprint arXiv2103.11733, 2021.
[12] R. van der Hofstad, *Preferential attachment models*, in: Random Graphs and Complex Networks (Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 2016, pp. 256–300.
[13] E. Jacob and P. Mörters, *Robustness of scale-free spatial networks*, Ann. Probab. **45** (2017), 1680–1722.
[14] L. Lüchtrath, *Percolation in weight-dependent random connection models*, PhD Thesis, University of Cologne, 2022.
[15] P. Mörters and D. Preiss, *Tangent measure distributions of fractal measures*, Math. Annalen **312** (1998), 53–99.
[16] M. D. Penrose and J. E. Yukich, *Weak laws of large numbers in geometric probability*, Ann. Appl. Probab. **13** (2003), 277–303.
[17] D. Preiss, *Geometry of measures in* $\mathbf{R}^n$*: distribution, rectifiability, and densities*, Ann. of Math. (2) **125** (1987), 537–643.
[18] S. Resnick, *Extreme Values, Regular Variation, and Point Processes*, Springer, 1987.
[19] U. Zähle, *Self-similar random measures I*, Prob. Th. Rel. Fields **80** (1988), 79–100.

Manuscript received October 14 2022
revised January 30 2023

P. Mörters
Universität zu Köln, Germany
E-mail address: `moerters@math.uni-koeln.de`