THE EVOLUTION OF RANDOM GRAPHS
ON SURFACES OF NON-CONSTANT GENUS

C. DOWDEN, M. KANG, M. MOSSHAMMER AND P. SPRÜSSEL

ABSTRACT. Given a graph G, the *genus* of G denotes the smallest integer g for which G can be drawn on the orientable surface of genus g without crossing edges. For integers $g, m \geq 0$ and $n > 0$, we let $S_g(n, m)$ denote the graph taken uniformly at random from the set of all graphs on $\{1, 2, \ldots, n\}$ with exactly $m = m(n)$ edges and with genus at most $g = g(n)$. We investigate the evolution of $S_g(n, m)$ as m increases, focussing on the number $|H_1|$ of vertices in the largest component. For $g \ll n$, we show that $|H_1|$ exhibits two phase transitions, one at around $m = \frac{n}{2}$ and a second one at around $m = n$. The exact behaviour of $|H_1|$ in the critical windows of these phase transitions depends on the order of $g = g(n)$.

1. BACKGROUND AND MOTIVATION

The theory of random graphs has a rich history, with many interesting models and numerous exciting results. Among the most fascinating discoveries has been the existence of thresholds at which dramatic changes occur in the fundamental properties of the graph, such as the typical order of the largest component or the probability of being connected.

Many of the features of the classical Erdős-Rényi random graph $G(n, m)$, taken uniformly at random from the set of all graphs on $[n] := \{1, 2, \ldots, n\}$ with exactly m edges, are now known. Prominent among these results has been the study of a phase transition in the 'evolution' of $G(n, m)$ around $m = \frac{n}{2}$, where the graph typically changes in structure from being planar with relatively small components of logarithmic order to being non-planar with a 'giant' component of linear order.

Various other interesting models of random graphs have also been investigated, and a particular focus has been the study of random graphs on surfaces [**1, 2, 4, 5, 6, 8, 9, 10, 11**]. The most popular case here is the random planar graph $P(n, m)$, taken uniformly at random from the set of all planar graphs on $[n]$ with exactly m edges.

It is well known that $G(n, m)$ is itself planar whp[1] if $m < \frac{n}{2} - \omega(n^{2/3})$, and so $P(n, m)$ thus behaves in the same way as $G(n, m)$ for this range of m (see e.g. [**7**]).

[1]With high probability, meaning with probability tending to 1 as $n \to \infty$.

However, different properties have been found to emerge when m is beyond this region [**2, 5, 6, 8**]. In particular, it was shown by Kang and Łuczak [**8**] that there are, perhaps rather surprisingly, two critical periods for $P(n, m)$. The first of these occurs when the giant component is formed, which happens at $m = \frac{n}{2} + O(n^{2/3})$ and thus coincides with when $G(n, m)$ starts to become non-planar. The second is then when the giant component covers nearly all vertices, which happens at $m = n + O(n^{3/5})$ and does not correspond to a phase transition in $G(n, m)$.

An interesting and natural generalisation of $P(n, m)$ is to consider random graphs on other surfaces. Here, we define an orientable surface of genus g as a sphere to which g handles have been attached, and the genus of a graph as the smallest value of g for which the graph can be embedded on such a surface without any crossing edges. We then define the graph $S_g(n, m)$ as being taken uniformly at random from the set of all graphs on $[n]$ with exactly m edges and genus at most g. Note that we thus have $S_0(n, m) = P(n, m)$.

In existing results on $S_g(n, m)$, only the constant genus model has been studied, and the actual value of g seems to have little impact. However, the genus can certainly have substantial influence if it is instead allowed to grow as a function of n, since it is clear that for sufficiently large $g = g(n)$ (e.g. $g = \binom{n}{2}$), the graph $S_g(n, m)$ is identical to the standard Erdős-Rényi random graph $G(n, m)$. Hence, $S_g(n, m)$ in fact generalises both $P(n, m)$ and $G(n, m)$, depending on the value of g, and so the study of $S_g(n, m)$ for general g appears to have the potential to be the most interesting and exciting area of all.

2. Results

The main contribution of this work is to analyse the order of the largest component of $S_g(n, m)$ for general genus $g = g(n)$, concentrating on the most interesting region of $m = m(n)$ from $m = \frac{n}{2} + \omega\left(n^{2/3}\right)$ (when $G(n, m)$ starts to become non-planar whp) to $m = (1 + o(1))n$ (when the second critical period in $P(n, m)$ occurs).

Note that $S_g(n, m)$ will be indistinguishable from $G(n, m)$ for this region of m if $g = \omega(n)$, since $G(n, m)$ can clearly only have genus at most $m = \Theta(n)$ (see [**3**]). However, in the first of our main results, we rather excitingly discover that for the weakly supercritical regime (i.e. when $m = \frac{n}{2} + s$ for positive s satisfying $n^{2/3} \ll s \ll n$) there are in fact *three* possible cases: the largest component of $S_g(n, m)$ can behave as with $P(n, m)$ (if g grows 'slowly'); part-way between $P(n, m)$ and $G(n, m)$ (if g grows at 'intermediate' speed); or as with $G(n, m)$ (if g grows 'fast').

To formalise these speeds of g, let us introduce the notion of contiguity.

Definition 1. We say that two random graph models $A(n)$ and $B(n)$ are *contiguous* if, for all properties $\mathcal{P}(n)$, it is the case that $A(n)$ has property $\mathcal{P}(n)$ whp if and only if $B(n)$ has property $\mathcal{P}(n)$ whp. We then say that $T = T(n, m)$ is a contiguity threshold for $S_g(n, m)$ and $G(n, m)$ if, for all fixed $\epsilon > 0$, $S_g(n, m)$ and $G(n, m)$ are contiguous for $g \geq (1 + \epsilon)T$ and are not contiguous for $g \leq (1 - \epsilon)T$.

Our three speeds will correspond to the cases (i) $g = o(T)$ (slow), (ii) $g = (c + o(1))T$ for $c \in (0, 1)$ (intermediate speed), and (iii) $g \geq (1 - o(1))T$ (fast). Precise results for T were derived in [**3**, Theorem 6.4]. In particular, when $m = \frac{n}{2} + s$ for $s > 0$ satisfying $n^{2/3} \ll s \ll n$, then $T = \frac{8s^3}{3n^2}$, and when $m \sim \lambda n$ for $\lambda > \frac{1}{2}$, then $T = \mu(\lambda)m$ for a given function $\mu(\lambda)$.

Given a random variable $a(n)$ and a positive-valued function $b(n)$, we shall find it convenient to adapt the standard deterministic asymptotic order notation into whp order notation by always replacing 'for all large n' in the standard definitions with 'whp'. For example, $a(n) = o(b(n))$ whp or $a(n) \ll b(n)$ whp means that, given any constant $\epsilon > 0$, we have $\frac{|a(n)|}{b(n)} < \epsilon$ whp.

For brevity, we shall also henceforth use $H_1 = H_1(S_g(n, m))$ to denote the largest component of $S_g(n, m)$, and $|H_1|$ to denote its order (i.e. its number of vertices). Then we are now ready to state our first main result.

Theorem 2. *Let* $m = m(n) = \frac{n}{2} + s \in \mathbb{N}$ *for* $s = s(n) > 0$ *satisfying* $n^{2/3} \ll s \ll n$, *and recall that* $T = T(n, m) = \frac{8s^3}{3n^2}$ *denotes the contiguity threshold for* $S_g(n, m)$ *and* $G(n, m)$. *Furthermore, let* $g = g(n)$ *and let* $H_1 = H_1(S_g(n, m))$. *Then whp*

$$
\begin{array}{rll}
\text{(i)} & |H_1| = (2 + o(1))s & \text{if } g = o(T), \\
\text{(ii)} & |H_1| = (f(c) + o(1))s & \text{if } g = (c + o(1))T \text{ for fixed } c \in (0, 1), \\
\text{and} \quad \text{(iii)} & |H_1| = (4 + o(1))s & \text{if } g \geq (1 + o(1))T,
\end{array}
$$

where $f\colon (0, 1) \to (2, 4)$ *is a fixed, concave, monotonically increasing, continuous function satisfying* $f(c) \to 2$ *as* $c \to 0$ *and* $f(c) \to 4$ *as* $c \to 1$.

Observe that the $g = o(T)$ region includes $P(n, m)$, while the $g \geq (1 + o(1))T$ region includes $G(n, m)$. Furthermore, $|H_1|$ lies on a 'sliding scale' between these two extremes if g grows at an intermediate speed.

Note in particular that if $1 \ll g \ll n$, then all three types of behaviour will occur as $m = \frac{n}{2} + s$ moves through the weakly supercritical regime. For example, if $g = n^{1/2}$, then $|H_1|$ behaves as in $G(n, m)$ for $s \leq (1 + o(1)) \left(\frac{3}{8}\right)^{1/3} n^{5/6}$, then part-way between $G(n, m)$ and $P(n, m)$ when $s \sim c_2 n^{5/6}$ for $c_2 > \left(\frac{3}{8}\right)^{1/3}$, and then as in $P(n, m)$ for $s \gg n^{5/6}$.

Our next result concerns the region of m from $m = \left(\frac{1}{2} + \epsilon\right) n$ for small $\epsilon > 0$ up to $m = (1 - \epsilon)n$. Recall that, for this region, $S_g(n, m)$ and $G(n, m)$ will certainly be contiguous if $g = \omega(n)$, and will not be contiguous if $g = o(n)$.

When $m \sim \lambda n$ for $\lambda > \frac{1}{2}$, the order of the largest component of $G(n, m)$ is known to be whp $(1 + o(1))\rho n$, where $\rho(\lambda) \in (0, 1)$ is uniquely determined by the equation $1 - \rho = e^{-2\lambda\rho}$ (and hence ρ grows smoothly, converging to 1 as $\lambda \to \infty$, see e.g. [**7**]). Note in particular that the number of vertices outside the largest component is thus whp $\Theta(n)$ when $m = \Theta(n)$.

For $P(n, m)$, by contrast, there exists a critical period at $m = (1 + o(1))n$, where the number of vertices outside the giant component drops to $o(n)$ whp. This occurs because in the intermediate region, when $m \sim \lambda n$ for $\lambda \in \left(\frac{1}{2}, 1\right)$, the

order of the largest component of $P(n, m)$ is whp $(2\lambda - 1 + o(1))n$. For $g = o(n)$, we find that $|H_1|$ behaves as in the planar model.

Theorem 3. *Let $m = m(n) \in \mathbb{N}$ satisfy $\left(\frac{1}{2} + \epsilon\right) n \leq m \leq (1 - \epsilon)n$ for some fixed $\epsilon > 0$, let $g = g(n) = o(n)$, and let $H_1 = H_1(S_g(n, m))$. Then whp*

$$|H_1| = (1 + o(1))(2m - n).$$

For this $g = o(n)$ case, we then investigate in detail the critical period at $m = (1 + o(1))n$ (for technical reasons, our precise upper bound will be $n + O\left((\log n)^{-2/3}n\right)$). We focus on the order of the graph $R = R(S_g(n, m)) := S_g(n, m) \backslash H_1$ outside the largest component, discovering fascinating and surprising behaviour.

In the planar $g = 0$ case, for $m = n + t$ we have $|R| = (2 + o(1))|t|$ when $t < 0$ and $n^{3/5} \ll |t| \ll n$, then $|R| = \Theta\left(n^{3/5}\right)$ when $t = O\left(n^{3/5}\right)$, and finally $|R| = \left(\left(\frac{n}{t}\right)^{3/2}\right)$ when $t > 0$ and $n^{3/5} \ll t \ll n$ (see [**8**]). For $S_g(n, m)$, we obtain an intriguing generalisation of this result.

Theorem 4. *Let $m = m(n) = n + t \in \mathbb{N}$ for $t = t(n)$ satisfying $t = o(n)$ if $t < 0$ and $t = O\left((\log n)^{-2/3}n\right)$ if $t > 0$, let $g = g(n) = o(n)$, let $\bar{g} = \bar{g}(n) = \max\left\{g, n^{3/5}\right\}$, and let $R = R(S_g(n, m))$. Then whp*

$$|R| = \begin{cases} (2 + o(1))|t| & \text{if } t < 0 \text{ and } |t| = \omega(\bar{g}), \\ \Theta(\bar{g}) & \text{if } t = o(\bar{g}), \\ \Theta\left(\left(\frac{n}{t}\right)^{3/2}\right) & \text{if } t > 0 \text{ and } t = \omega(\bar{g}). \end{cases}$$

One particularly startling phenomenon to note is the sudden drop in the value of $|R|$ immediately after the $t = o(\bar{g})$ region when $g = \omega\left(n^{3/5}\right)$. For instance, if $g = n^{0.9}$, then we have $|R| = \Theta\left(n^{0.9}\right)$ whp when $t = n^{0.9 - \epsilon}$, but $|R| = \Theta\left(n^{0.15 - \frac{3\epsilon}{2}}\right)$ whp when $t = n^{0.9 + \epsilon}$. The transition from order $\Theta\left(n^{0.9}\right)$ to order $\Theta\left(n^{0.15}\right)$ must take place in the range $t = c\bar{g}$, where $c > 0$. By contrast, the behaviour of $|R|$ seems to be smooth leading up to the $t = o(\bar{g})$ region.

Observe also that the phase transition shown in Theorem 4 is similar to the planar case if $g = O\left(n^{3/5}\right)$. However, in contrast to Theorems 2 and 3, there are differences to the planar case as soon as g reaches $\omega\left(n^{3/5}\right)$, even though this is still substantially below the $T = \Theta(n)$ contiguity threshold.

3. Techniques

Our methodology draws on many areas, and involves new advancements and developments. Several different tools are consequently used, including probabilistic, enumerative, and analytic techniques.

One important approach is via the core-kernel decomposition. Here, a graph is (a) split into complex and non-complex components, (b) each complex component is decomposed into a 2-core and a forest, and then (c) the 2-core is decomposed into its kernel and a subdivision of edges. By reversing these instructions (starting

from an 'appropriate' kernel) and keeping track of the number of possible ways to perform the various steps, information on the typical structure and properties of the constructed graph can hence be obtained.

Note that the core-kernel approach involves expressing a set of graphs as a sum of subclasses involving the relevant parts of the composition (such as the 2-core and the kernel). We may then attempt to determine the main contribution to such a sum by using a combinatorial variant of Laplace's method from real analysis.

Another technique that we utilise is the (seemingly simple, but very useful) idea of double-counting. For instance, suppose that we wish to relate the number of graphs in two sets $\mathcal{G}_n$ and $\mathcal{G}'_n$. For each graph in $\mathcal{G}_n$, we would aim to construct many graphs in $\mathcal{G}'_n$ by making various alterations (e.g. adding/deleting edges in suitable places), and we would then try to show that each graph in $\mathcal{G}'_n$ is not constructed too many times.

The challenge when creating such a proof lies in finding a successful way to construct many graphs of the desired type without introducing a large amount of double-counting. Hence, the alterations used in the construction process need to be carefully controlled, in order to allow some way of bounding the number of possibilities for the original graph.

Finally, let us mention that one other useful way to gain knowledge of $S_g(n, m)$ is by applying recent results from [3] on the genus of the Erdős-Rényi random graph $G(n, m)$. Note that if $G(n, m)$ is known to have genus at most $g(n)$ whp, for some specified values of $g(n)$ and $m(n)$, then this immediately implies that for these values the random graphs $G(n, m)$ and $S_g(n, m)$ must be contiguous.

References

1. Chapuy G., Fusy É., Giménez O., Mohar B. and Noy M., *Asymptotic enumeration and limit laws for graphs of fixed genus*, J. Combin. Theory Ser. A **118** (2011), 748–777.
2. Dowden C., *The evolution of uniform random planar graphs*, Electron. J. Combin. **17** (2010), #7.
3. Dowden C., Kang M. and Krivelevich M., *The genus of the Erdős-Rényi random graph and the fragile genus property*, Random Structures Algorithms (2019), to appear, `https://doi.org/10.1002/rsa.20871`.
4. Dowden C., Kang M. and Sprüssel P., *The evolution of random graphs on surfaces*, SIAM J. Discrete Math. **32** (2018), 695–727.
5. Gerke S., McDiarmid C., Steger A. and Weißl A., *Random planar graphs with n nodes and a fixed number of edges*, in: Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '05), 2005, 999–1007.
6. Giménez O. and Noy M., *Asymptotic enumeration and limit laws of planar graphs*, J. Amer. Math. Soc. **22** (2009), 309–329.
7. Janson S., Łuczak T. and Ruciński A., *Random Graphs*, Wiley-Interscience, New York, 2000.
8. Kang M. and Łuczak T., *Two critical periods in the evolution of random planar graphs*, Trans. Amer. Math. Soc. **364** (2012), 4239–4265.
9. Kang M., Moßhammer M. and Sprüssel P., *Phase transitions in graphs on orientable surfaces*, `arXiv:1708.07671`, submitted for publication.
10. McDiarmid C., *Random graphs on surfaces*, J. Combin. Theory Ser. B **98** (2008), 778–797.
11. McDiarmid C., Steger A. and Welsh D., *Random planar graphs*, J. Combin. Theory Ser. B **93** (2005), 187–205.

C. Dowden, Graz University of Technology, Austria,
e-mail: `dowden@math.tugraz.at`

M. Kang, Graz University of Technology, Austria,
e-mail: `kang@math.tugraz.at`

M. Moßhammer, Graz University of Technology, Austria,
e-mail: `mosshammer@math.tugraz.at`

P. Sprüssel, Graz University of Technology, Austria,
e-mail: `spruessel@math.tugraz.at`