University of Wrocław

Faculty of Mathematics and Computer Science

Mathematical Institute

speciality: actuarial and financial mathematics

Uniwersytet Wrocławski

Wydział Matematyki i Informatyki

Instytut Matematyczny

specjalność: matematyka aktuarialno-finansowa

Natalia Jeszka

Gromov-Hausdorff on a circle

Gromov-Hausdorff na okręgu

Master's thesis
written under the supervision of
dr Piotr Dyszewski

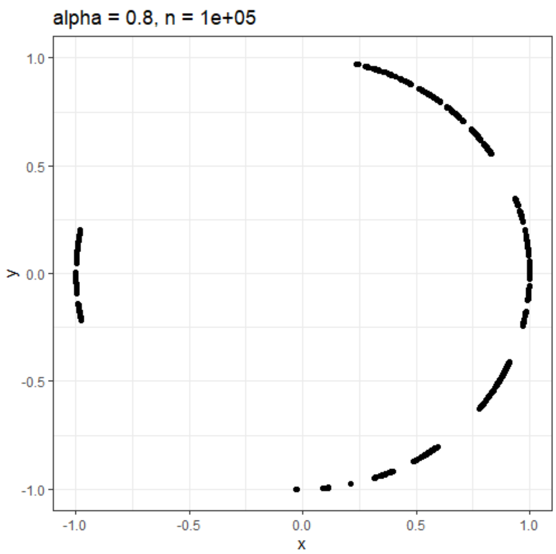
Praca magisterska
napisana pod kierunkiem
dr. Piotra Dyszewskiego

1 Introduction

The complexity of some random structures, especially the most general ones, may effectively prevent a deeper understanding of certain characteristics of considered object. Even though this difficulty might be partially overcome e.g. by introducing a second level of randomness (like a stochastic process), we are going to realize that the most reliable way to the complete comprehension of the analysed structure is the simplification.

In our case the mentioned random but simplified structure will be a graph with fixed outline, namely a cycle, so the randomness will not come from uncertain neighbourhoods of vertices but from the random distances between any two neighbours. Thinking of random distances, it naturally comes up by itself to consider random metric spaces which the graph might be perfectly well. The machinery provided a.o. by Burago, Burago and Ivanov in [2] allows then to, having already a sequence of such random metric spaces, formulate a limit theorem which introduces a fine approximation of the graphs.

Our main goal will be proving exactly such a theorem, and we are going to achieve it in a couple of varying cases, depending on the considered random metric. One might suspect that with number of vertices growing larger and larger (and with proper scaling of the distances between the vertices), a cycle graph is going to have a shape of a circle. Indeed, it will appear so – but only under assumption that distance between any two vertices won't be significantly larger than the distance between any other two. What will giving up on this assumption change? These notably large distances which will become allowed are going to produce random "gaps" in the circle so the limit object will be a random sum of arcs. This structure we may imagine as



1.1 The model under study

Let us reckon a sequence of i.i.d. random variables $(Y_i)_{i\in[n]}$ such that $Y_i > 0$ and where, as usually, $[n] = \{1, 2, \dots, n\}$. (For now we don't make further assumptions on the distribution of Y_i, as we would like to consider various cases in the future.)

We assign Y_i as weights to the edges of a cyclic graph on n vertices such that edge $(k, k+1)$ has weight Y_k for $k = 1, \dots, n-1$, and edge $(n, 1)$ has weight Y_n. To this weighted graph we will refer to as $\mathcal{C}_n$. It will be treated as a metric space and when we do not define the metric explicitly we understand that it is given by weights Y_i, to be precise: by the formula

$$d_{\mathcal{C}_n}(1, k+1) = \min\left(\sum_{i=1}^{k} Y_i,\ \sum_{i=k+1}^{n} Y_i\right)$$

which in a moment is going to become clearer. Some cases require normalization of these weights and then the metric will be given.

1.2 First passage percolation

In the light of our further interests in the limit behaviour of $\mathcal{C}_n$, the study of first passage percolation on the graph might be considered a side comment. However, it provides much deeper understanding of the structure of $\mathcal{C}_n$ as well as the natural metric on this graph. Thanks to that, it may serve as an interesting introduction and become an ingathering of, henceforward, very useful facts.

The percolation model is supposed to imitate the progressive exploration of the graph in which the cost of traveling between the adjacent vertices is given by the weight of the appropriate edge. The distance between any two vertices is then the minimal weight of a path connecting them i.e.

$$d(i, j) = \min_{\pi: i\to j} \sum_{e\in\pi} Y_e.$$

For now, we would like to consider only weights with $\mathbb{E}Y_i = m < \infty$. In the far future, after this case is already well understood, we will give up on this assumption, letting $m = \infty$.

The concept of moving on the cycle is now therefore quite simple as any two vertices i, j are connected by exactly two paths (of lengths $|i - j|$ and $n - |i - j|$) so, in order to formalize our fist observations, it seems convenient to denote the length of the path from vertex 1 to $k + 1$ by

$$S_k = \sum_{i=1}^{k} Y_i$$

and the distance on $\mathcal{C}_n$ by

$$d_{\mathcal{C}_n}(1, k+1) = \min(S_k,\ S_n - S_k).$$

Moreover, as the weights are i.i.d., it should be probable that the above value is obtained for the shorter of the two possible paths. It indeed turns out to be so, which can be put in the following

Lemma 1.1. Let $\epsilon > 0$ and $m < \infty$. For sequence $\{k_n\}_n$ such that $n(\epsilon + \frac{1}{2}) \leqslant k_n \leqslant n$ for every n, we have

$$\mathbb{P}[S_{k_n} < S'_{n-k_n}] \xrightarrow{0} .$$

Above $S'_{n-k_n} = \sum_{i=1}^{n-k_n} Y'_i$ for some $Y'_i \overset{d}{=} Y_i$.

Before proving the lemma let us note that it actually formalizes prior intuitions. First of all, by the choice of k_n, path corresponding with S_{k_n} is the longer one (in the sense of having more elements). What is more, if H_n denotes the distance between vertex 1 and another randomly chosen vertex of $\mathcal{C}_n$, and U_n has uniform distribution on $[n]$, then

$$H_n \overset{d}{=} S_{U_n} \wedge S'_{n-U_n}.$$

That is due to the fact that $\mathbb{P}[H_n < t] = \mathbb{P}[S_{U_n} < t \vee S_{n-U_n} < t] = \mathbb{P}[S_{U_n} \wedge S_{n-U_n} < t]$. The lemma claims that the probability of obtaining this minimum above for the sum consisting of more elements, approaches 0 in the limit. Therefore it indeed convinces that the optimal path on the cycle is given by the shorter arc.

Proof. For the proof we rearrange

$$\mathbb{P}[S_{k_n} < S'_{n-k_n}] = \mathbb{P}\Big[\sum_{i=1}^{k_n} Y_i \; - \sum_{i=1}^{n-k_n} Y'_i < 0\Big] = \mathbb{P}\Big[\sum_{i=1}^{n-k_n} Y_i - Y'_i \; + \sum_{i=n-k_n+1}^{k_n} Y_i < 0\Big].$$

Basing on the law of large numbers and the choice of k_n, we would like to approximate

$$\sum_{i=n-k_n+1}^{k_n} Y_i \approx (2k_n - n)m \geqslant 2\epsilon nm > 0$$

and

$$\sum_{i=1}^{n-k_n} Y'_i - Y_i = o(n)$$

because $\mathbb{E}[Y'_i - Y_i] = 0$. That would explain vanishing of the probability above:

$$\mathbb{P}[S_{k_n} < S'_{n-k_n}] \approx \mathbb{P}[2\epsilon nm + 0 < 0] = 0.$$

To do this formally, let us split the considered event

$$\Big\{\sum_{i=1}^{n-k_n} Y_i - Y'_i \; + \sum_{i=n-k_n+1}^{k_n} Y_i < 0\Big\}$$

into a disjoint union to obtain three addends converging (by the law of large numbers) to 0 and addend:

$$\mathbb{P}\Big[\sum_{i=1}^{n-k_n} Y_i - Y'_i \; + \sum_{i=n-k_n+1}^{k_n} Y_i < 0, \; \Big|\frac{\sum_{i=1}^{n-k_n} Y_i - Y'_i}{n - k_n}\Big| < \epsilon^2, \; \Big|\frac{\sum_{i=n-k_n+1}^{k_n} Y_i}{2k_n - n} - m\Big| < \epsilon^2\Big] \leqslant$$
$$\mathbb{P}[-\epsilon^2(n - k_n) + m(2k_n - n) - \epsilon^2(2k_n - n) < 0] = \mathbb{P}[m(2k_n - n) < \epsilon^2 k_n] \leqslant$$
$$\mathbb{P}[2mn\epsilon < \epsilon^2 n] = \mathbb{P}[2m < \epsilon].$$

Because ϵ was arbitrary, so without loss of generality $\epsilon < m/2$, it altogether assures that

$$\mathbb{P}\Big[\sum_{i=1}^{n-k_n} Y_i - Y_i' \; + \sum_{i=n-k_n+1}^{k_n} Y_i < 0\Big] \xrightarrow{n} 0.$$

□

Now that we are certain which is the optimal path on the cycle and therefore – what is the proper way of measuring distances, we would also like to ask what is usual distance between two randomly chosen vertices (when n is getting large). It turns out that the answer to that question is already within our reach and can be specified in the explicit form of the limit distribution.

However, for the clarity, we will precede by proving a few simple proprieties of convergence in probability and distribution which may already be familiar from the probability theory courses. The first of those lemmas assures that if a sequence of random variables converges almost surely, we may only look on some random (but diverging to $+\infty$ in probability) indices and observe the same limit.

Lemma 1.2. Let Z_n, A_n and A be random variables such that $Z_n \xrightarrow{\mathbb{P}} \infty$ and $A_n \xrightarrow{a.s.} A$ as $n \to \infty$. In such case

$$A_{Z_n} \xrightarrow{\mathbb{P}} A.$$

Proof. Let us fix $\delta, M > 0$, and as often while dealing with convergence in probability, split

$$\mathbb{P}[|A_{Z_n} - A| > \delta] = \mathbb{P}[|A_{Z_n} - A| > \delta, Z_n \geqslant M] + \mathbb{P}[|A_{Z_n} - A| > \delta, Z_n < M].$$

We assumed that $\mathbb{P}[Z_n \geqslant M] \xrightarrow{1}$ with $n \to \infty$ and therefore

$$\mathbb{P}[|A_{Z_n} - A| > \delta, Z_n < M] \xrightarrow{0}.$$

As for the first component of the sum, from the definition of the almost sure convergence

$$1 = \mathbb{P}[\exists M > 0 \; \forall n > M \; |A_n - A| < \delta] = \lim_{M \to \infty} \mathbb{P}[\forall n > M \; |A_n - A| < \delta],$$

where last equality holds by the continuity of measure as the sequence of events

$$\big\{ \{\forall n > M \; |A_n - A| < \delta\} \big\}_{M \in \mathbb{N}}$$

is ascending. Meanwhile

$$0 \xleftarrow{M \to \infty} \mathbb{P}[\exists n > M \; |A_n - A| > \delta] \geqslant \mathbb{P}[|A_{Z_n} - A| > \delta, Z_n \geqslant M],$$

and the claim follows. □

The lemma above is needed as the justification of the following expansion of the law of large numbers.

Corollary 1.3. For S_n, U_n and $m = \mathbb{E}Y_i$ defined previously we have

$$\frac{S_{U_n}}{U_n} \xrightarrow{\mathbb{P}} m.$$

Proof. Clearly $U_n \xrightarrow{\mathbb{P}} \infty$, and from the strong law of large numbers $\frac{S_n}{n} \xrightarrow{a.s.} m$. Applying Lemma 1.2 secures the claim. □

Next step leading to the understanding of the mean distance between random vertices in the cycle is another simple exercise saying that if we uniformly choose one integer from the interval of length n and divide by this length, in fact we choose one of the equidistant points in the interval $(0,1]$. Increasing n refines the partition, so for n large we actually pick a point from $(0,1)$.

Proposition 1.4. For the random variable $\frac{U_n}{n}$ the limit distribution is uniform on $(0,1)$.

Proof. For the justification, it's enough to compute the limit of the distribution function:

$$\mathbb{P}\Big[\frac{U_n}{n} \le t\Big] = \mathbb{P}\big[U_n \le nt\big] = \frac{\lfloor nt \rfloor}{n} \xrightarrow{n} t = \mathbb{P}\big[U \le t\big].$$

□

We are now approaching, already announced, main statement of this section which describes the limit law of the mean distance between two points on the cycle. Of course, it confirms all the presumptions suggesting that, having one vertex fixed, it's enough to choose the second one uniformly and measure the shorter arc.

Theorem 1.5. Let $V \sim \mathcal{U}(0,1)$ and $m < \infty$. Then

$$\frac{H_n}{n} \Rightarrow (V \wedge (1-V))m.$$

Proof. The convergence above follows easily from our initial analysis of the $\mathcal{C}_n$ properties. An elegant way of concluding it uses the continuous mapping theorem.

To that end, let us first notice that the function $f(x,y,z) = (xy) \wedge (z(1-x))$ is clearly continuous as the composition of continuous functions. Further on, observe that considered random variable is of the form

$$\frac{H_n}{n} \overset{d}{=} \left(\frac{S_{U_n}}{U_n} \cdot \frac{U_n}{n}\right) \wedge \left(\frac{S_{n-U_n}}{n-U_n} \cdot \frac{n-U_n}{n}\right).$$

Now, as $n-U_n \sim \mathcal{U}[n-1]$, Lemma 1.2 gives the convergence $\frac{S_{n-U_n}}{n-U_n} \xrightarrow{\mathbb{P}} m$ while the Corollary 1.3 provides $\frac{S_{U_n}}{U_n} \xrightarrow{\mathbb{P}} m$. In the Proposition 1.4 we just noticed that $\frac{U_n}{n} \Rightarrow V$. To arrive at the conclusion, it is enough to combine the above observations and apply continuous mapping theorem:

$$\left(\frac{S_{U_n}}{U_n} \cdot \frac{U_n}{n}\right) \wedge \left(\frac{S_{n-U_n}}{n-U_n} \cdot \frac{n-U_n}{n}\right) \Rightarrow (mV) \wedge (m(1-V)).$$

□

2 Gromov-Hausdorff distance and other technicalities

Having already introduced the space, a graph on n vertices, whose limit we'll be looking for while n will be getting larger, we are still in need of the language which allows us to speak about the limit of metric spaces at all. It means that we would like to be able to measure distances between metric spaces and, as a consequence, say that two metric spaces are "close" (or "far") form each other.

As it turns out, the concept of Hausdorff and Gromov-Hausdorff distance is going to be very useful here. We shall now, following [2], introduce both definitions and restate some of the facts and properties applicable for us in the future.

Wherever W, X, X_n or Z will be appearing below, they always denote arbitrary metric spaces with metrics d_W, d_X, d_{X_n} and d_Z respectively.

Definition 2.1. Suppose W and X are both subspaces of Z. We define the Hausdorff distance between W and X via

$$d_H(W,X) = \inf\{r > 0 : X \subseteq U_r(W),\ W \subseteq U_r(X)\},$$

where $U_r(S)$ denotes the r-neighbourhood of the set S i.e. $U_r(S) = \{x : d_Z(x,S) < r\}$.

We may actually think of the Hausdorff distance as of the longest distance from some point of W to any point of X (in the sense of metric on Z) and the other way around, i.e.

$$d_H(W,X) = \max\left\{\sup_{x\in X} d_Z(x,W),\ \sup_{w\in W} d_Z(w,X)\right\},$$

however, as the previous formula is going to be more convenient for us, we keep it an original definition.

Let us note now that, even though the Hausdorff metric might seem very natural way of measuring distance between metric spaces, it strongly requires the existence of the superspace Z. We would like to be able to give up on this assumption and define distinct metrics on W and X. For this purpose, form the Hausdorff we derive the Gromov-Hausdorff metric in the most intuitive way: we look for the closest isometric embeddings of W and X into the common superspace Z and take the Hausdorff distance (which for such embeddings is already well defined):

Definition 2.2. The Gromov-Hausdorff distance between spaces W and X is given via

$$d_{GH}(W,X) = \inf\{r > 0 : (\exists W' \cong W,\ X' \cong X,\ Z \supseteq W', X')\ d_H(W',X') = r\}.$$

Above the mark $\cong$ denotes the isometric isomorphism between the spaces.
Moreover, the sequence X_n converges in the Gromov-Hausdorff sense to X when

$$\lim_{n\to\infty} d_{GH}(X_n, X) = 0.$$

Remark 2.3. Considering any metric spaces $X' \cong X$ and $X'_n \cong X_n$ such that $X', X'_n \subseteq Z$, directly from the definition of Gromov-Hausdorff distance we have $d_{GH}(X_n, X) \leqslant d_H(X'_n, X')$. As a consequence, convergence in Hausdorff sense may imply convergence in Gromov-Hausdorff sense, i.e. if $d_H(X'_n, X') \xrightarrow{0}$, then also $d_{GH}(X_n, X) \xrightarrow{0}$.

With the tool for measuring how close the two metric spaces are from each other, let us now also provide a handful of theory which in some cases makes the computations of the Gromov-Hausdorff distance much easier. It will concern a simple concept of correspondence relation.

Definition 2.4. The correspondence between sets W and X is the set $\mathcal{R} \subseteq W \times X$ such that

$$\forall w \in W \ \exists x \in X \ (w, x) \in \mathcal{R} \text{ and } \forall x \in X \ \exists w \in W \ (w, x) \in \mathcal{R}.$$

The above definition refers to the relation in which every element of W has some related element of X and symmetrically: every element of X is related to some element of W. Having this in mind, we come across the simplest possible example: for any surjection $f : W \to X$, we can define a correspondence as

$$\mathcal{R} = \{(w, f(w)) : w \in W\}.$$

This form is so well-known that it is often refered to as the correspondence associated with function f.

All the correspondences that we will consider on the following pages will be associated with some surjections, therefore let us not dwell upon other forms of correspondences but investigate its possible usefulness. Due to that, we need to introduce one more concept:

Definition 2.5. Given the correspondence $\mathcal{R} \subseteq W \times X$, we compute its distortion as

$$\operatorname{dis}(\mathcal{R}) = \sup\{|d_W(w, w') - d_X(x, x')| : (w, x), (w', x') \in \mathcal{R}\}.$$

When $\mathcal{R}$ is associated with $f : W \to X$, the formula above may be simplified:

$$\operatorname{dis}(\mathcal{R}) = \sup_{w, w' \in W} |d_W(w, w') - d_X(f(w), f(w'))|.$$

Remark 2.6. Obvious but important observation to be made here is that if $\mathcal{R}$ is associated with an isometry from W to X, then $d_W(w, w') = d_X(f(w), f(w'))$ for all w, w' which means that $\operatorname{dis}(\mathcal{R}) = 0$. (In fact, the opposite implication is also true, i.e. if $\operatorname{dis}(\mathcal{R}) = 0$, then $\mathcal{R}$ is associated with some isometry – but this property would not be that useful for us.)

When it comes to usefulness, we can easily see the utility of the above comment after we get familiar with the following

Lemma 2.7. We can express the Gromov-Hausdorff distance between W and X as

$$d_{GH}(W,X) = \frac{1}{2} \inf_{\mathcal{R}_{W\leftrightarrow X}} (\mathrm{dis}(\mathcal{R}))$$

where the infimum is taken over all correspondences $\mathcal{R}_{W\leftrightarrow X}$ between W and X. It might be, in other words, stated as

$$d_{GH}(W,X) = \inf\{r > 0 : \exists \mathcal{R}_{W\leftrightarrow X}\ \mathrm{dis}(\mathcal{R}) < 2r\}.$$

The proof basis on the triangle inequality and, even though it is not very complicated one, we shall provide just a brief idea and for the details refer to Bugaro et al. [2].

Sketch of proof. Without loss of generality assume that $W, X \subseteq Z$ for some Z, and take $r > 0$ such that $d_H(W,X) < r$. Then

$$\mathcal{R} = \{(w,x) :\ w \in W,\ x \in X,\ d_Z(w,x) < r\}$$

is a well-defined correspondence. Triangle inequality assures $\mathrm{dis}(\mathcal{R}) < 2r$ which gives us ($\leqslant$) inequality from the claim.

For the remaining ($\geqslant$) inequality, consider arbitrary correspondence $\mathcal{R}$ with $\mathrm{dis}(\mathcal{R}) = 2r$. It's enough to metrize $W \cup X$ in such way that

$$d_{W\cup X}|_{W\times W} = d_W,\ d_{W\cup X}|_{X\times X} = d_X \text{ and } d_H(W,X) < r \text{ in } (W \cup X, d_{W\cup X}).$$

It can be obtained by setting $d_{W\cup X}(w,w') = d_W(w,w')$, $d_{W\cup X}(x,x') = d_X(x,x')$ and

$$d_{W\cup X}(w,x) = \inf\{d_W(w,w') + r + d_X(x,x') :\ (w',x') \in \mathcal{R}\}$$

for $w, w' \in W$ and $x, x' \in X$. The first two conditions are then obviously satisfied. The last one can be checked as well, and the idea behind it is that we've defined $d_{W\cup X}$ such that it sets distance between w and x to be equal to r when $(w,x) \in \mathcal{R}$ and close to r otherwise. □

We conclude that when dealing with Gromov-Hausdorff distance directly from definition begins to seem overwhelming, it's enough to define such correspondences between spaces X_n and X that their distortions tend to 0 with $n \to \infty$. By the above lemma, that would presently mean that X_n converge to X in the Gromov-Hausdorff sense.

3 Case of finite mean

Our main interest now will be examination of the "shape" of $\mathcal{C}_n$ when $n \to \infty$. At the beginning we will stay with the $\mathbb{E}Y_i = m < \infty$ but next section will serve as its supplement with $\mathbb{E}Y_i = \infty$.

As it was already announced, one can suspect that the more vertices we join in the cycle and zoom out, the more it will look like a circle. However, this intuition seems to be true only in the finite mean case. When mean begins to be infinite, there appears a possibility that, even when there are plenty of vertices, a space between two consecutive ones happens that large that after zooming out we observe "a gap" in the limit circle.

To avoid these gaps in the finite mean case, we need to make sure that presence of large edges is quite improbable – and this fact, provided exactly by the finiteness of the mean, will be the contents of the next lemma.

Lemma 3.1. In the considered case $m < \infty$ we observe

$$n\mathbb{P}\big[Y_i > n\big] \xrightarrow{n} 0.$$

Proof. Let us first note that $\mathbb{1}_{\{Y_i>n\}} \leqslant \frac{\mathbb{1}_{\{Y_i>n\}}Y_i}{n}$ and thus

$$\mathbb{P}[Y_i > n] = \mathbb{E}[\mathbb{1}_{\{Y_i>n\}}] \leqslant \mathbb{E}\Big[\frac{\mathbb{1}_{\{Y_i>n\}}Y_i}{n}\Big] = \frac{\mathbb{E}[\mathbb{1}_{\{Y_i>n\}}Y_i]}{n}.$$

So $n\mathbb{P}[Y_i > n] \leqslant \mathbb{E}[\mathbb{1}_{\{Y_i>n\}}Y_i]$ and by the integrability of Y_i we may use dominated convergence theorem to write

$$\lim_{n\to\infty} \mathbb{E}[\mathbb{1}_{\{Y_i>n\}}Y_i] = \mathbb{E}[Y_i \lim_{n\to\infty} \mathbb{1}_{\{Y_i>n\}}] = 0.$$

□

Remark 3.2. Following the proof we make sure that lemma remains true after dividing Y_i by any constant $c > 0$, i.e. in the form $n\mathbb{P}\big[Y_i > nc\big] \xrightarrow{n} 0$.

This simple fact has a natural, equally simple consequence: if a large weight appears in the graph with very low probability, then even the largest present edge should seem insignificant with comparison to the remaining ones. It will be the contents of the next lemma, for which needs we will denote

$$M_n = \max_{i\in[n]} Y_i$$

and stay with this notation later on, if necessary.

Lemma 3.3. With all of the introduced notation, in the case $m < \infty$ we have

$$\mathbb{P}[2M_n > S_n] \xrightarrow{n} 0.$$

Perhaps the statement is clearer when rewritten as: $\mathbb{P}[M_n > S_n - M_n] \xrightarrow{0}$. Here we see that lemma investigates the probability of one edge in the graph being greater than the sum of all other edges – so the probability of event that going from some vertex to one of the neighbours is optimal via all remaining vertices, not directly via shared edge. Of course, as we already discussed, to make it probable, there would also need to be a chance that in the cycle there is at least one large weight. However, as the previous lemma states, probability of such event is approaching 0 in limit, therefore the probability of having an edge larger than whole remaining arc will behave alike.

Proof. Similarly to the proof of Lemma 1.1, we would like to aproximate S_n by $n\mathbb{E}Y_i$. Again, for a fixed $\epsilon > 0$, we do it by splitting:

$$\mathbb{P}[2M_n > S_n] = \mathbb{P}[2M_n > S_n,\ |\frac{S_n}{n} - m| > \epsilon] + \mathbb{P}[2M_n > S_n,\ |\frac{S_n}{n} - m| \leqslant \epsilon].$$

By the law of large numbers, we should only care about the second component because the first disappears in the limit: $\mathbb{P}[2M_n > S_n,\ |\frac{S_n}{n} - m| > \epsilon] \xrightarrow{n} 0$.

And when it comes to mentioned second part:

$$\mathbb{P}[2M_n > S_n,\ |\frac{S_n}{n} - m| \leqslant \epsilon] = \mathbb{P}[2M_n > S_n,\ n(m+\epsilon) \geqslant S_n \geqslant n(m-\epsilon)] \leqslant$$
$$\mathbb{P}[2M_n > S_n,\ S_n \geqslant n(m-\epsilon)] \leqslant \mathbb{P}[2M_n > n(m-\epsilon)] = \mathbb{P}[M_n > \frac{n(m-\epsilon)}{2}] \leqslant$$
$$n\mathbb{P}[Y_i > \frac{n(m-\epsilon)}{2}] \xrightarrow{n} 0.$$

Above the convergence follows from Lemma 3.1 and the last inequality is a consequence of the simple observation that

$$\mathbb{P}[M_n > a] \leqslant \mathbb{P}[Y_1 > a] + \mathbb{P}[Y_2 > a] + \cdots + \mathbb{P}[Y_n > a] = n\mathbb{P}[Y_i > a]$$

for any a (as the maximum has to be obtained for some of Y_i so in fact $\mathbb{P}[M_n > a] = \mathbb{P}[Y_i > a]$ for some i). ☐

Remark 3.4. Just like in case of the previous lemma, we can notice that a constant here doesn't really play a role. Having constant 2 in the formula of the lemma provides nice, already discussed interpretation. However, following the proof, we notice that 2 can be replaced with any $c > 0$ and the statement stays true.

As we have already previewed in the introduction, the natural metric on the cycle comes, of course, from the percolation – so is given by the weights of the edges. Let us recall that we denoted it by $d_{\mathcal{C}_n}$. So for example, $d_{\mathcal{C}_n}(1,k) = S_{k-1} \wedge (S_n - S_{k-1})$.

Despite the fact that $d_{\mathcal{C}_n}$ is here the most natural metric, it leads to the graph of (random) girth S_n and thus if the limit object was circle, it should be of circumference S_n. We would like to normalize this metric such that the radius of the limit circle becomes equal to 1 (so both girth and circumference are 2π, so much less confusing than previously).

3.1 Stochastic normalization

First and probably the simplest idea of such a normalization is clearly by $2\pi/S_n$. Its main disadvantage is stochastic nature (which luckily, with some effort, can be removed; we are going to provide an argument for convergence also in this case).

However, as it's simpler, it allows to see desired convergence much more straightforward than its deterministic analogue. Thanks to that it will be very useful, especially because we are still in need of some examples of measuring the Gromov-Hausdorff distance. And this quite directly normalized metric $d_{\mathcal{C}_n}$ will allow us to vividly describe Gromov-Hausdorff distance between the cycle $\mathcal{C}_n$ and the circle $\mathcal{C} = \{x \in \mathbb{R}^2 : \|x\| = 1\}$.

Theorem 3.5. Consider $X_n = (\mathcal{C}_n, \frac{2\pi d_{\mathcal{C}_n}}{S_n})$ and $X = (\mathcal{C}, d_{\mathcal{C}})$, where $d_{\mathcal{C}}$ denotes the length of the shorter arc on the circle. For such metric spaces, as $n \to \infty$, we observe

$$X_n \xrightarrow{\mathbb{P}} X$$

in the Gromov-Hausdorff sense.

Proof. We would like to show that

$$d_{GH}(X_n, X) \xrightarrow{n} 0.$$

Having in mind Remark 2.3, it's enough to define metric spaces $X_n' \cong X_n$, $X' \cong X$ and $Z \supseteq X_n', X'$ such that $d_H(X_n', X') \xrightarrow{\mathbb{P}} 0$.

As the planar graph and the unit circle naturally live on the plane, we might want to metrize $\mathbb{R}^2$ such that it is possible to isometrically embed there both X_n and X. However, note that the only crucial aspect of this metric on $\mathbb{R}^2$ would be preserving distances on the unit circle (outside this circle we could define distances freely, just to keep the metric axioms true). Therefore, we realize that metrizing only the circle and embedding X_n and X there will work perfectly well for us.

Having that in mid, we consider $Z = X' = X = (\mathcal{C}, d_{\mathcal{C}})$ and below define isometry $\psi : \mathcal{C}_n \to \mathcal{C}$ for which we would have $d_H(\psi(X_n), X') \xrightarrow{\mathbb{P}} 0$ (and for which we will write $X_n' = \psi(X_n)$).

We would like ψ to map the vertices of $\mathcal{C}_n$ on the circle, preserving distances given by the normalized weights $\frac{2\pi Y_i}{S_n}$. It can be obtained simply by fixing one point on the circle, say $(0, 1)$, mapping vertex 1 on this point and, having that, mapping 2 on the point which is exactly $\frac{2\pi Y_1}{S_n}$ clockwise distant form 1 on the circle (so with respect to $d_{\mathcal{C}}$). Similarly we proceed for $3, \ldots, n-1$.

It may not be obvious at first glance that ψ is well defined, i.e. that after repeating this operation $n-1$ times, the remaining arc has length exactly $\frac{2\pi Y_n}{S_n}$. Luckily, it is in fact so because:

$$2\pi - \frac{2\pi}{S_n}\sum_{i=1}^{n-1} Y_i = 2\pi - \frac{2\pi}{S_n}(S_n - Y_n) = \frac{2\pi Y_n}{S_n}.$$

To prove required convergence, let us now fix any $\epsilon > 0$ and show that

$$\mathbb{P}[d_H(X'_n, X') \leqslant \epsilon] \xrightarrow{n} 1.$$

That means we would like to understand

$$\inf\{r > 0 : X'_n \subseteq U_r(X'),\ X' \subseteq U_r(X'_n)\}.$$

At the beginning, note that first condition is trivial because for any $r > 0$ we have $X'_n \subseteq X' \subseteq U_r(X')$ (let us recall: X'_n are points representing vertices of the graph mapped on the circle and X' is the full circle). This observation reduces formula above to the

$$\inf\{r > 0 : X' \subseteq U_r(X'_n)\}.$$

Translating, we ask how probable it is that, for very small r, taking r-neighbourhood of the points on circle, we would cover whole circle. And, as the intuition suggests, it turns out that when $n \to \infty$ so we have more and more of these points, it becomes very probable.

The neighbourhood $U_r(x)$ for some $x \in X_n$ is an arc of length $2r$ (r to the left of x and r to the right of x). Considering two consecutive points $x_1, x_2 \in X'_n$, we may ask: how large should r be to entirely cover the arc between x_1 and x_2 with $U_r(X'_n)$? Simply, it should be at least the half of the distance between those points. And now: considering all pairs of consecutive points, how large r do we need to cover all the arcs with $U_r(X'_n)$? Of course, it's enough to cover the longest arc (if for some r the neighbourhood $U_r(X'_n)$ covers the longest "gap" between points of X'_n, it clearly covers all shorter ones). Therefore

$$\inf\{r > 0 : X' \subseteq U_r(X'_n)\} = \frac{1}{2}\frac{2\pi M_n}{S_n}.$$

Summarizing, we end up with

$$\mathbb{P}[d_H(X'_n, Y') < \epsilon] = \mathbb{P}[\frac{\pi M_n}{S_n} \leqslant \epsilon] = \mathbb{P}[M_n \leqslant \frac{\epsilon}{\pi} S_n]$$

which looks quite familiar to the result of Lemma 3.3, just with different constant. Fortunately, we already noticed in the Remark 3.4 that this constant is not important and can be replaced. Thus the lemma provides

$$\mathbb{P}[M_n \leqslant \frac{\epsilon}{\pi} S_n] \xrightarrow{n} 1$$

which completes the case of stochastic normalization. □

3.2 Deterministic normalization

As previously announced, we would now like to put some effort in normalizing metric on $\mathcal{C}_n$ with deterministic constant. Originally the part responsible for randomness in the normalization was S_n which is, as we always hope, well approximated by nm. This approximation is the only change in the statement of the following theorem with comparison to Theorem 3.5, so the limit object remains the same and we can recall it to be the unit circle $\mathcal{C} = \{x \in \mathbb{R}^2 : ||x|| = 1\}$ with the shorter arc metric $d_{\mathcal{C}}$.

Theorem 3.6. Consider $X_n = (\mathcal{C}_n, \frac{2\pi d_{\mathcal{C}_n}}{nm})$ and $X = (\mathcal{C}, d_{\mathcal{C}})$. For such metric spaces, as $n \to \infty$, we have

$$X_n \xrightarrow{\mathbb{P}} X$$

in the Gromov-Hausdorff sense.

Even though the variation seems minor, it requires significant modification of the proof. Trying to follow the proof of the stochastic normalization case would fail in the moment of defining function ψ which in its primary form would not be an isometry here (e.g. because the girth of the $\mathcal{C}_n$ is not anymore 2π but instead $2\pi \pm \delta$ for some small δ). Keeping ψ as an isometry would require changing the distinguished circle for one with circumference equal to girth of $\mathcal{C}_n$ and that would also call for reconsideration of the form of X. To avoid these complications (and also: present another way to deal with Gromov-Hausdorff convergence), we reach for already introduced language of correspondences and distortions.

To recall the application of Lemma 2.7: if we succeed in defining such correspondences between spaces X_n and X that their distortions tend to 0 in probability, we actually show Gromov-Hausdorff convergence of X_n to X (also in probability), so we prove stated theorem – and this is our current intention.

We will associate our correspondence with a surjection $f : \mathcal{C} \to \mathcal{C}_n$ because it is the most natural way to define such a relation. And as distortion estimates how much differently corresponding spaces measure distances, we would like f to be "as close to isometry as possible".

Let us begin the construction of f with the mapping $\psi : \mathcal{C}_n \to \mathcal{C}$ used in the stochastic normalization case. We have already noticed that it is not an isometry, however it is close enough to isometry for us (the only distance which it may not preserve is between n and 1, and the deviation there is rather insignificant). We can imagine reversed operation: taking a point which lies on the circle and is associated with a vertex of the cycle and assigning this vertex as value at considered point of the circle. When it comes to not yet considered points of $\mathcal{C}$, we would like f to map them on the closest possible vertex of $\mathcal{C}_n$ (these vertices are already located on the circle, so it makes sense to say which is the closest with respect to $d_{\mathcal{C}}$). Slightly more formally:

Proof of Theorem 3.6. Let us divide the unit circle $\{x \in \mathbb{R}^2 : \|x\| = 1\}$ into n arcs of a random length such that the first one is $\frac{Y_1+Y_n}{S_n}\pi$ and the k-th one: $\frac{Y_k+Y_{k-1}}{S_n}\pi$ long (for $k = 2,\dots,n$). For the sake of clarity, we say that those arcs are open on the right and closed on the left when considered clockwise on the circle. The function $f : \mathcal{C} \to \mathcal{C}_n$ is then defined as

$$f(x) = k$$

when x lies in the k-th arc of the circle.

This function is obviously a surjection and therefore defines correspondence $\mathcal{R} = \{(x, f(x)) : x \in \mathcal{C}\}$ between the circle and the cycle. Its distortion is then computed as follows.

Consider $x, x' \in \mathcal{C}$. If both x and x' are in the same arc (let's say it is k-th arc), we have:

$$d_{\mathcal{C}_n}(f(x), f(x')) = d_{\mathcal{C}_n}(k,k) = 0$$

and

$$d_{\mathcal{C}}(x,x') \leqslant \frac{2M_n}{S_n}\pi$$

as this is the bound for the length of any arc.

Therefore, in this case, the probability that the supremum appearing in the definition of distortion will be large indeed disappears at infinity:

$$\mathbb{P}\left[\sup_{\substack{x,x'\in\mathcal{C} \\ \text{in the same arc}}} \left|d_{\mathcal{C}}(x,x') - d_{\mathcal{C}_n}(f(x), f(x'))\right| > \epsilon\right] =$$

$$\mathbb{P}\left[\sup_{\substack{x,x'\in\mathcal{C} \\ \text{in the same arc}}} d_{\mathcal{C}}(x,x') > \epsilon\right] \leqslant \mathbb{P}\left[\frac{2M_n}{S_n}\pi > \epsilon\right] = \mathbb{P}\left[M_n > \frac{\epsilon}{2\pi}S_n\right] \xrightarrow{n} 0$$

and final convergence is again the result of the Lemma 3.3 together with the Remark 3.4.

On the other hand, assume that x and x' lie in different arcs. For simplicity, let's suppose that x is in the first arc and x' in the k-th one. It's enough to consider only this case thanks to possibility to renumerate arcs. Then:

$$d_{\mathcal{C}_n}(f(x), f(x'))\frac{2\pi}{nm} = d_{\mathcal{C}_n}(1,k)\frac{2\pi}{nm} =$$

$$\left((Y_1 + \dots + Y_{k-1})\frac{2\pi}{nm}\right) \wedge \left((Y_k + \dots + Y_n)\frac{2\pi}{nm}\right) = \left(S_{k-1}\frac{2\pi}{nm}\right) \wedge \left((S_n - S_{k-1})\frac{2\pi}{nm}\right)$$

and

$$\left(\left(\frac{Y_1}{2} + Y_2 + \dots + Y_{k-2} + \frac{Y_{k-1}}{2}\right)\frac{2\pi}{S_n}\right) \wedge \left(\left(S_n - \left(\frac{Y_n}{2} + Y_1 + \dots + Y_{k-1} + \frac{Y_k}{2}\right)\right)\frac{2\pi}{S_n}\right)$$

$$\leqslant d_{\mathcal{C}}(x,x') \leqslant$$

$$\left(\left(\frac{Y_n}{2} + Y_1 + \dots + Y_{k-1} + \frac{Y_k}{2}\right)\frac{2\pi}{S_n}\right) \wedge \left(\left(S_n - \left(\frac{Y_1}{2} + Y_2 + \dots + Y_{k-2} + \frac{Y_{k-1}}{2}\right)\right)\frac{2\pi}{S_n}\right).$$

For clarification, imagine situations of placing x and x' first in the closer endpoints of arcs 1 and k, and second in the further endpoints of these arcs – in these cases we see that distance between considered points has to be larger that at least one of the arcs of lengths $\left(\frac{Y_1}{2} + Y_2 + \cdots + Y_{k-2} + \frac{Y_{k-1}}{2}\right)\frac{2\pi}{S_n}$ and $\left(S_n - \left(\frac{Y_n}{2} + Y_1 + \cdots + Y_{k-1} + \frac{Y_k}{2}\right)\right)\frac{2\pi}{S_n}$ and shorter then both arcs of lengths $\left(\frac{Y_n}{2} + Y_1 + \cdots + Y_{k-1} + \frac{Y_k}{2}\right)\frac{2\pi}{S_n}$ and $\left(S_n - \left(\frac{Y_1}{2} + Y_2 + \cdots + Y_{k-2} + \frac{Y_{k-1}}{2}\right)\right)\frac{2\pi}{S_n}$. It justifies the above bounds.

Examining those bounds, we observe that if we could only replace nm by S_n, then during subtraction (needed to compute distortion) a lot of addends would simplify and most probably we would be able to finish with calculation similar to the one performed in the first part of the proof. As this seems to be the main problem, again, we hope to engage the law of large numbers to find a solution. Indeed, this will soon be the case, however, let us first write explicitly:

$$\sup_{\substack{x,x'\in\mathcal{C} \\ x \text{ in 1st arc, } x' \text{ in } k\text{th arc}}} \left| d_{\mathcal{C}_n}(1,k)\frac{2\pi}{nm} - d_{\mathcal{C}}(x,x') \right| =$$

$$\left| \left(\left(\frac{Y_1}{2} + Y_2 + \cdots + Y_{k-2} + \frac{Y_{k-1}}{2} \right) \wedge \left(S_n - \left(\frac{Y_n}{2} + Y_1 + \cdots + Y_{k-1} + \frac{Y_k}{2} \right) \right) \right) \frac{2\pi}{S_n} - (S_{k-1} \wedge (S_n - S_{k-1})) \frac{2\pi}{nm} \right| \vee$$

$$\left| \left(\left(\frac{Y_n}{2} + Y_1 + \cdots + Y_{k-1} + \frac{Y_k}{2} \right) \wedge \left(S_n - \left(\frac{Y_1}{2} + Y_2 + \cdots + Y_{k-2} + \frac{Y_{k-1}}{2} \right) \right) \right) \frac{2\pi}{S_n} - (S_{k-1} \wedge (S_n - S_{k-1})) \frac{2\pi}{nm} \right|.$$

We shall quickly simplify this expression by noting that, in probability, when $S_{k-1} < S_n - S_{k-1}$ then also

$$\frac{Y_1}{2} + Y_2 + \cdots + Y_{k-2} + \frac{Y_{k-1}}{2} < S_n - \left(\frac{Y_n}{2} + Y_1 + \cdots + Y_{k-1} + \frac{Y_k}{2} \right)$$

and

$$\frac{Y_n}{2} + Y_1 + \cdots + Y_{k-1} + \frac{Y_k}{2} < S_n - \left(\frac{Y_1}{2} + Y_2 + \cdots + Y_{k-2} + \frac{Y_{k-1}}{2} \right).$$

Therefore

$$\left| \left(\frac{Y_1}{2} + Y_2 + \cdots + Y_{k-2} + \frac{Y_{k-1}}{2} \right) \frac{2\pi}{S_n} - S_{k-1}\frac{2\pi}{nm} \right| \vee$$
$$\left| \left(\frac{Y_n}{2} + Y_1 + \cdots + Y_{k-1} + \frac{Y_k}{2} \right) \frac{2\pi}{S_n} - S_{k-1}\frac{2\pi}{nm} \right| \vee$$
$$\left| \left(S_n - \left(\frac{Y_n}{2} + Y_1 + \cdots + Y_{k-1} + \frac{Y_k}{2} \right) \right) \frac{2\pi}{S_n} - (S_n - S_{k-1})\frac{2\pi}{nm} \right| \vee$$
$$\left| \left(S_n - \left(\frac{Y_1}{2} + Y_2 + \cdots + Y_{k-2} + \frac{Y_{k-1}}{2} \right) \right) \frac{2\pi}{S_n} - (S_n - S_{k-1})\frac{2\pi}{nm} \right|.$$

We will argue that all of the differences converge to 0 in probability. Let us begin with rearranging the first one:

$$S_{k-1}\frac{2\pi}{nm}-\left(\frac{Y_1}{2}+Y_2+\cdots+Y_{k-2}+\frac{Y_{k-1}}{2}\right)\frac{2\pi}{S_n}=$$
$$S_{k-1}\left(\frac{2\pi}{nm}-\frac{2\pi}{S_n}\right)+\frac{2\pi}{S_n}\left(S_{k-1}-\left(\frac{Y_1}{2}+Y_2+\cdots+Y_{k-2}+\frac{Y_{k-1}}{2}\right)\right)=$$
$$\frac{S_{k-1}2\pi(S_n-nm)}{nmS_n}+\frac{2\pi}{S_n}\left(\frac{Y_1}{2}+\frac{Y_{k-1}}{2}\right).$$

We again treat both components separately. The first disappears almost surely because $S_{k-1}<S_n$ and due to the strong law of large numbers:

$$\frac{S_{k-1}2\pi(S_n-nm)}{nmS_n}=\frac{2\pi}{m}\cdot\frac{S_{k-1}}{S_n}\left(\frac{S_n}{n}-m\right)\leqslant\frac{2\pi}{m}\left(\frac{S_n}{n}-m\right)\xrightarrow{a.s.}0.$$

The second can be bounded:

$$\frac{2\pi}{S_n}\left(\frac{Y_1}{2}+\frac{Y_{k-1}}{2}\right)\leqslant\frac{2\pi M_n}{S_n}$$

and by Lemma 3.3 and Remark 3.4 we know that the above fraction is arbitrary small with probability approaching 1 when $n\to\infty$.

Likewise, for the second difference in the maximum representing value of distortion:

$$\left(\frac{Y_n}{2}+Y_1+\cdots+Y_{k-1}+\frac{Y_k}{2}\right)\frac{2\pi}{S_n}-S_{k-1}\frac{2\pi}{nm}=$$
$$\left(\frac{Y_n}{2}+Y_1+\cdots+Y_{k-1}+\frac{Y_k}{2}-S_{k-1}\right)\frac{2\pi}{S_n}+S_{k-1}\left(\frac{2\pi}{S_n}-\frac{2\pi}{nm}\right)=$$
$$\left(\frac{Y_n}{2}+\frac{Y_k}{2}\right)\frac{2\pi}{S_n}+\frac{2\pi}{m}\cdot\frac{S_{k-1}}{S_n}\left(m-\frac{S_n}{n}\right)\leqslant\frac{2\pi M_n}{S_n}+\frac{2\pi}{m}\left(m-\frac{S_n}{n}\right)\xrightarrow{\mathbb{P}}0,$$

and for the third:

$$(S_n-S_{k-1})\frac{2\pi}{nm}-\left(S_n-\left(\frac{Y_n}{2}+Y_1+\cdots+Y_{k-1}+\frac{Y_k}{2}\right)\right)\frac{2\pi}{S_n}=$$
$$(S_n-S_{k-1})\left(\frac{2\pi}{nm}-\frac{2\pi}{S_n}\right)+\left((S_n-S_{k-1})-S_n+\left(\frac{Y_n}{2}+Y_1+\cdots+Y_{k-1}+\frac{Y_k}{2}\right)\right)\frac{2\pi}{S_n}=$$
$$\frac{2\pi}{m}\cdot\frac{S_n-S_{k-1}}{S_n}\left(\frac{S_n}{n}-m\right)+\left(\frac{Y_n}{2}+\frac{Y_k}{2}\right)\frac{2\pi}{S_n}\leqslant\frac{2\pi}{m}\left(\frac{S_n}{n}-m\right)+\frac{2\pi M_n}{S_n}\xrightarrow{\mathbb{P}}0,$$

and for the fourth:

$$\left(S_n-\left(\frac{Y_1}{2}+Y_2+\cdots+Y_{k-2}+\frac{Y_{k-1}}{2}\right)\right)\frac{2\pi}{S_n}-(S_n-S_{k-1})\frac{2\pi}{nm}=$$
$$\left(S_n-\left(\frac{Y_1}{2}+Y_2+\cdots+Y_{k-2}+\frac{Y_{k-1}}{2}\right)-(S_n-S_{k-1})\right)\frac{2\pi}{S_n}+(S_n-S_{k-1})\left(\frac{2\pi}{S_n}-\frac{2\pi}{nm}\right)=$$
$$\frac{2\pi}{S_n}\left(\frac{Y_1}{2}+\frac{Y_{k-1}}{2}\right)+\frac{2\pi}{m}\cdot\frac{S_n-S_{k-1}}{S_n}\left(m-\frac{S_n}{n}\right)\leqslant\frac{2\pi M_n}{S_n}+\frac{2\pi}{m}\left(m-\frac{S_n}{n}\right)\xrightarrow{\mathbb{P}}0.$$

Summarizing, $\operatorname{dis}(\mathcal{R})\xrightarrow{\mathbb{P}}0$ and therefore $(\mathcal{C}_n,\frac{2\pi d_{\mathcal{C}_n}}{nm})\xrightarrow{\mathbb{P}}(\mathcal{C},d_{\mathcal{C}})$ in the Gromov-Hausdorff sense. □

4 Case of infinite mean

Now as we already know the limit object in the case of finite mean, we will also find out that $\mathcal{C}_n$ behaves analogously when expectation of the edges' weights is infinite. We consider their distribution given via:

$$\mathbb{P}[Y_i > x] = x^{-\alpha}$$

for some $\alpha \in (0,1)$ and all $x \geqslant 1$.

It clearly contradicts the assumptions of Lemma 3.1. Also, the intuitive difference is that presently we can actually expect few huge edge weights among several significantly smaller. All of that clearly means that the Lemma 3.3 as well doesn't have to be true now.

Despite the lack of these tools, we still are able to somehow understand the distribution of S_n. It involves some theory of α-stable laws, which we now recall after Durrett [3].

4.1 α-stable laws

We are going to motivate our reasoning by the underneath theorem.

Theorem 4.1. Let $Y_1, Y_2, \dots$ be i.i.d. random variables such that:

1. $\lim_{x\to\infty} \mathbb{P}[Y_1 > x]/\mathbb{P}[|Y_1| > x] = \theta \in [0,1]$ and
2. $\mathbb{P}[|Y_1| > x] = x^{-\alpha} l(x)$

for some $\alpha < 2$ and some function l satisfying $\lim_{x\to\infty} l(tx)/l(x) = 1$ for all $t > 0$. Then

$$\frac{\sum_{i=1}^n Y_i - b_n}{a_n} \Rightarrow L \qquad \text{as } n \to \infty$$

where L has non-degenerate distribution and

$$a_n = \inf\{x : \mathbb{P}[|Y_1| > x] \leqslant 1/n\}, \qquad b_n = n\mathbb{E}[Y_1 \mathbb{1}_{\{|Y_1| \leqslant a_n\}}].$$

Such distributions, i.e. distributions of L above form, are usually called *stable* or *α-stable*. Formally, L is said to have a stable law when $L \overset{d}{=} \frac{\sum_{i=1}^k L_i - d_k}{c_k}$ for some constants c_k, d_k and sequence of i.i.d. random variables $L_i \overset{d}{=} L$, for every k. However, it is a consequence of another theorem that one can unambiguously characterise stable laws by limit laws of $\frac{\sum_{i=1}^k Y_i - d_k}{c_k}$ for some sequence of i.i.d. Y_i.

Let us now make two quick comments on the statement of the theorem. First one is purely nomenclatural but useful while working with the literature on the subject, second actually simplifies the claim of the theorem for our needs.

Remarks 4.2. 1. Function l fulfilling mentioned condition $\lim_{x\to\infty} l(tx)/l(x) = 1$ for all $t > 0$ is often called *slowly varying*.

2. It can be shown that if $\alpha < 1$ (which is true in our case), we can let $b_n = 0$ and the statement of the theorem remains true.

We are now going to reinterpret the foregoing theorem for our needs. At the beginning, notice that both assumed conditions are trivial in our case as our variables Y_i are non-negative, so:

1. $\lim_{x\to\infty} \mathbb{P}[Y_1 > x]/\mathbb{P}[|Y_1| > x] = \lim_{x\to\infty} \mathbb{P}[Y_1 > x]/\mathbb{P}[Y_1 > x] = 1$ and
2. $\mathbb{P}[|Y_1| > x] = \mathbb{P}[Y_1 > x] = x^{-\alpha} \cdot 1$ and $l(x) = 1$ is slowly varying (likewise every constant).

Like we already remarked, because we consider only $\alpha \in (0,1)$, we can forget about the centring b_n in the statement of the theorem. Let us compute scaling constant:

$$a_n = \inf\{x : \mathbb{P}[Y_1 > x] \leqslant \frac{1}{n}\} = \inf\{x : x^{-\alpha} \leqslant n^{-1}\} = \inf\{x : x \geqslant n^{1/\alpha}\} = n^{1/\alpha}.$$

As we were denoting $\sum_{i=1}^{n} Y_i = S_n$, all the preceding observations justify the following

Corollary 4.3. In the case of $\mathbb{P}[Y_i > x] = x^{-\alpha}$ for $\alpha \in (0,1)$ and $x \geqslant 1$ we have

$$n^{-1/\alpha} S_n \Rightarrow L \qquad \text{as } n \to \infty$$

for some α-stable L.

Unfortunately, described convergence isn't good enough for us as it only allows to follow S_n as a sequence. For the technical reasons it would be more convenient to be able to consider a stochastic process defined for $t \in [0,1]$ but somehow reflecting behaviour of $\{S_k\}_{k\leqslant n}$. The definition of such a process is quite natural, we simply write

$$L_n(t) = n^{-1/\alpha} S_{[nt]}.$$

Having succeeded in defining it for all $t \in [0,1]$, we would like to, from now on, work in the Skorokhod space $\mathcal{D}[0,1]$ of *càdlàg* functions on $[0,1]$ (i.e. functions continuous on the right with limits on the left). It seems reasonable to consider such a space because of the following

Claim 4.4. The processes $(L_n(t))_{t\in[0,1]}$ defined as above are random elements of $\mathcal{D}[0,1]$. It is the direct result of the observation that

$$L_n(t) = \sum_{k=0}^{n-1} n^{-1/\alpha} S_k \mathbb{1}_{[\frac{k}{n}, \frac{k+1}{n})}(t).$$

The connection between processes L_n and considered graph is now straightforward. In fact, of the vertices of $\mathcal{C}_n$ we can think as of the points on $[0,1]$, scaled later to $[0,2\pi]$ and "wrapped in circle", i.e.:

$$\mathcal{G}_n = \left\{ e^{2\pi i \frac{L_n(t)}{L_n(1)}} \;:\; t \in [0,1] \right\}.$$

Claim 4.5. Indeed, it can be formalized by noting that $(\mathcal{G}_n, d_{\mathcal{C}}) \cong (\mathcal{C}_n, \frac{2\pi d_{\mathcal{C}_n}}{S_n})$ as the mapping

$$e^{2\pi i \frac{L_n(t)}{L_n(1)}} \mapsto k$$

for $t \in [\frac{k-1}{n}, \frac{k}{n})$ is a well-defined isometry.

Being able to describe $\mathcal{C}_n$ in terms of L_n, we suppose that understanding the limit of processes L_n will be useful in order to understand the limit of $\mathcal{C}_n$ as well. Therefore let us reach out to Resnick [5] for the following:

Theorem 4.6. Let $\{X_{n,j} : j \geqslant 1\}$ be a sequence of i.i.d. random vectors such that

$$n\mathbb{P}[X_{n,1} \in \cdot] \xrightarrow{\nu} (\cdot)$$

for some Lévy measure ν, and in the sense of vague convergence of measures. Assume also that

$$\lim_{\epsilon \to 0} \limsup_{n \to \infty} n\mathbb{E}[(X_{n,1})^2 \mathbb{1}_{\{|X_{n,1}| \leqslant \epsilon\}}] = 0.$$

Then the process defined via

$$X_n(t) = \sum_{k=1}^{[nt]} (X_{n,k} - \mathbb{E}[X_{n,k} \mathbb{1}_{\{|X_{n,k}| \leqslant 1\}}])$$

converges weakly to a Lévy jump process with Lévy measure ν. Mentioned convergence holds in the sense of convergence in the space of all *càdlàg* functions on $[0,1]$ (usually denoted by $\mathcal{D}[0,1]$) equipped with the J_1 metric which will be introduced at the beginning of the following subsection.

But first, to adopt this theorem for our needs, let us consider

$$X_{n,k} = n^{-1/\alpha} Y_k$$

and check that such a sequence satisfies the assumptions of the theorem.

As we consider the distribution $\mathbb{P}[X_{n,1} > t] = \mathbb{P}[n^{-1/\alpha} Y_1 > t] = \mathbb{P}[Y_1 > tn^{1/\alpha}] = t^{-\alpha} n^{-1}$, it follows that

$$n\mathbb{P}[X_{n,1} > t] = t^{-\alpha} = \nu(t, \infty)$$

which is a Lévy measure, i.e.

$$\int_{\mathbb{R} \setminus \{0\}} (|y|^2 \wedge 1) \nu(dy) < \infty.$$

Indeed,

$$\int_{\mathbb{R} \setminus \{0\}} (|y|^2 \wedge 1) \nu(dy) = \int_{(0,\infty)} (y^2 \wedge 1) \nu(dy) = \int_0^1 y^2 \nu(dy) + \int_1^\infty 1 \nu(dy)$$

and both integrals are finite as:

$$\int_1^\infty 1 \nu(dy) = \nu(1, \infty) = 1^{-\alpha} = 1$$

and by Radon-Nikodym theorem

$$\int_0^1 y^2 d\nu = \int_0^1 y^2 \frac{d\nu}{d\lambda} d\lambda = \int_0^1 y^2 \alpha y^{-\alpha-1} d\lambda = \alpha \int_0^1 y^{-\alpha+1} dy = \frac{\alpha}{2-\alpha}.$$

Also, the second assumption is true because:

$$n\mathbb{E}[(X_{n,1})^2 \mathbb{1}_{\{|X_{n,1}|\leqslant\epsilon\}}] = n\mathbb{E}[n^{-2/\alpha} Y_1^2 \mathbb{1}_{\{|n^{-1/\alpha}Y_1|\leqslant\epsilon\}}] = n^{1-\frac{2}{\alpha}} \mathbb{E}[Y_1^2 \mathbb{1}_{\{Y_1 \leqslant \epsilon n^{1/\alpha}\}}] =$$

$$n^{1-\frac{2}{\alpha}} \int_0^\infty \mathbb{P}[Y_1^2 \mathbb{1}_{\{Y_1\leqslant\epsilon n^{1/\alpha}\}} > t] dt = n^{1-\frac{2}{\alpha}} \int_0^{\epsilon^2 n^{\frac{2}{\alpha}}} \mathbb{P}[\epsilon n^{\frac{1}{\alpha}} \geqslant Y_1 > \sqrt{t}] dt =$$

$$n^{1-\frac{2}{\alpha}} \int_0^{\epsilon^2 n^{\frac{2}{\alpha}}} \mathbb{P}[Y_1 > \sqrt{t}] - \mathbb{P}[Y_1 > \epsilon n^{\frac{1}{\alpha}}] dt = n^{1-\frac{2}{\alpha}} \left(\int_1^{\epsilon^2 n^{\frac{2}{\alpha}}} t^{-\alpha/2} dt - \left(\epsilon^2 n^{\frac{2}{\alpha}} - 1\right) \left(\epsilon n^{\frac{1}{\alpha}}\right)^{-\alpha} \right) =$$

$$n^{1-\frac{2}{\alpha}} \left(\frac{(\epsilon^2 n^{\frac{2}{\alpha}})^{1-\frac{\alpha}{2}} - 1}{1-\frac{\alpha}{2}} - \epsilon^{2-\alpha} n^{\frac{2}{\alpha}-1} + \epsilon^{-\alpha} n^{-1} \right) = \frac{\epsilon^{2-\alpha} - n^{1-\frac{2}{\alpha}}}{1-\frac{\alpha}{2}} - \epsilon^{2-\alpha} + \epsilon^{-\alpha} n^{-\frac{2}{\alpha}}$$

and therefore $\lim_{\epsilon\to 0} \limsup_{n\to\infty} n\mathbb{E}[(X_{n,1})^2 \mathbb{1}_{\{|X_{n,1}|\leqslant\epsilon\}}] = 0$.

To apply the theorem we shall compute also the term $\mathbb{E}[X_{n,k} \mathbb{1}_{\{|X_{n,k}|\leqslant 1\}}]$ appearing in the claim. It is equal to

$$\mathbb{E}[X_{n,k} \mathbb{1}_{\{|X_{n,k}|\leqslant 1\}}] = n^{-\frac{1}{\alpha}} \mathbb{E}[Y_k \mathbb{1}_{\{Y_k \leqslant n^{\frac{1}{\alpha}}\}}] = n^{-\frac{1}{\alpha}} \int_0^{n^{\frac{1}{\alpha}}} \mathbb{P}[Y_k > t] dt = n^{-\frac{1}{\alpha}} \int_1^{n^{\frac{1}{\alpha}}} t^{-\alpha} dt =$$

$$n^{-\frac{1}{\alpha}} \frac{1}{1-\alpha} \left((n^{\frac{1}{\alpha}})^{1-\alpha} - 1 \right) = \frac{1}{n(1-\alpha)} - \frac{1}{n^{\frac{1}{\alpha}}(1-\alpha)}.$$

Therefore defining

$$X_n(t) = \sum_{k=1}^{[nt]} (X_{n,k} - \mathbb{E}[X_{n,k} \mathbb{1}_{\{|X_{n,k}|\leqslant 1\}}]) = \sum_{k=1}^{[nt]} \left(n^{-1/\alpha} Y_k - \frac{1}{n(1-\alpha)} + \frac{1}{n^{\frac{1}{\alpha}}(1-\alpha)} \right) =$$

$$L_n(t) - \frac{[nt]}{n(1-\alpha)} + \frac{[nt]}{n^{\frac{1}{\alpha}}(1-\alpha)},$$

by the above theorem we know that such processes have a weak limit. Denote it by $X_n(t) \Rightarrow X(t)$ and notice that

$$L_n(t) = X_n(t) + \frac{[nt]}{n(1-\alpha)} - \frac{[nt]}{n^{\frac{1}{\alpha}}(1-\alpha)} \Rightarrow X(t) + \frac{t}{1-\alpha}.$$

It justifies that L_n indeed converges weekly in $\mathcal{D}[0,1]$ and we can denote its limit by L. As X is Lévy process with Lévy measure ν so with characteristic function $\mathbb{E}\left[e^{isX(t)}\right] = e^{t\psi(s)}$ for

$$\psi(s) = \int_1^\infty \left(e^{isx} - 1\right) \alpha x^{-\alpha-1} dx + \int_0^1 \left(e^{isx} - 1 - isx\right) \alpha x^{-\alpha-1} dx,$$

therefore L is Lévy process with characteristic function $\mathbb{E}\left[e^{isL(t)}\right] = e^{t\left(\psi(s) + \frac{is}{1-\alpha}\right)}$.

4.2 Stochastic normalization

We already know that processes L_n are elements of the Skorokhod space $\mathcal{D}[0,1]$ and that they converge in distribution to L. By the merit of the Skorokhod's representation theorem, it means almost sure convergence in the sense of the metric on $\mathcal{D}[0,1]$ for some version of L_n and L. This metric we now introduce after Skorokhod [6].

Definition 4.7. Let $\mathcal{D}[0,1]$ be the space of all *càdlàg* functions on [0,1] and Λ be the family of increasing, continuous surjections from [0,1] to [0,1]. For $x, y \in \mathcal{D}[0,1]$ we define the distance by

$$d_{J_1}(x,y) = \inf_{\lambda \in \Lambda} \{ ||\lambda - I||_\infty \vee ||x - y \circ \lambda||_\infty \}$$

where I stands for identity mapping and $||\cdot||_\infty$ denotes the supremum norm.

In the light of the above definition, it would be convenient to understand the convergence of L_n in $\mathcal{D}[0,1]$ as almost sure convergence of the mentioned versions of these processes. It means the existence of $\lambda_n \in \Lambda$ such that

$$\lambda_n \rightrightarrows I \text{ and } L_n \circ \lambda_n \rightrightarrows L.$$

Let us recall that we were interested in finding the limit object for $\mathcal{G}_n = \{ e^{2\pi i \frac{L_n(t)}{L_n(1)}} : t \in [0,1] \}$. Knowing that $L_n \Rightarrow L$, we would now want to infer that also

$$\frac{L_n(\cdot)}{L_n(1)} \Rightarrow \frac{L(\cdot)}{L(1)}.$$

It is the result of the fact that $L_n(1) \xrightarrow{\mathbb{P}} L(1)$ (in $\mathbb{R}$) which can be deduced from the two following lemmas.

Lemma 4.8. Let f_n, f be *càdlàg* on $[0,1]$ and $x \in [0,1]$ be the continuity point of f. If $f_n \to f$ in $\mathcal{D}[0,1]$, then also $f_n(x) \to f(x)$ in $\mathbb{R}$.

Proof. Take $\lambda_n \in \Lambda$ indicating the convergence $f_n \to f$ in $\mathcal{D}[0,1]$, i.e. such that

$$\sup_{t \in [0,1]} |\lambda_n t - t| < \epsilon/2 \quad \text{and} \quad \sup_{t \in [0,1]} |f_n(t) - f(\lambda_n t)| < \epsilon/2$$

for any ϵ.

Now, respectively by: the triangle inequality, convergence $f_n \to f$ in $\mathcal{D}[0,1]$ and continuity of f at x we have:

$$|f_n(x) - f(x)| \leqslant |f_n(x) - f(\lambda_n x)| + |f(\lambda_n x) - f(x)| < \epsilon/2 + \epsilon/2 = \epsilon.$$

□

Claim 4.9. Lévy process L is continuous in probability at any point $t \in [0,1]$ so considering any sequence $t_n \to t$ we have $L(t_n) \xrightarrow{\mathbb{P}} L(t)$ (for the justification, see Resnick [5]).

It now clearly confirms supposition that $L_n(1) \xrightarrow{\mathbb{P}} L(1)$ in $\mathbb{R}$, and we are now ready to prove required weak convergence, i.e. we can state:

Lemma 4.10. Consider processes L_n defined above and their limit Lévy process L. Then "normalized" processes also converge weakly to the "normalized" limit:

$$\frac{L_n(\cdot)}{L_n(1)} \Rightarrow \frac{L(\cdot)}{L(1)}.$$

Proof. Take $\lambda_n \rightrightarrows I$ such that $L_n \circ \lambda_n \rightrightarrows L$ and for $t \in [0,1]$ write

$$\left|\frac{L_n(\lambda_n(t))}{L_n(1)} - \frac{L(t)}{L(1)}\right| \leqslant \left|\frac{L_n(\lambda_n(t))}{L_n(1)} - \frac{L(t)}{L_n(1)}\right| + \left|\frac{L(t)}{L_n(1)} - \frac{L(t)}{L(1)}\right| =$$
$$\frac{|L_n(\lambda_n(t)) - L(t)|}{L_n(1)} + \frac{L(t)\,|L(1) - L_n(1)|}{L_n(1)L(1)} \leqslant \frac{|L_n(\lambda_n(t)) - L(t)|}{L_n(1)} + \frac{|L(1) - L_n(1)|}{L_n(1)} \xrightarrow{n} 0$$

by the uniform convergence of $L_n \circ \lambda_n$, the convergence of $L_n(1)$ and the fact that, by the definition, $L_n(1) = n^{-1/\alpha} S_n > 0$ almost surely. ☐

We are also already able to spot the property which the limit process inherits from the sequence L_n:

Lemma 4.11. Process L is non-decreasing in probability.

Proof. Begin with observing that processes L_n are non-decreasing by definition (as increasing sums of non-negative random variables). It means that $L_n(t) - L_n(s) \geqslant 0$ for $s < t$. In such case

$$L(t) - L(s) \geqslant (L(t) - L_n(t)) + (L_n(s) - L(s)) \xrightarrow{n} 0$$

because of just claimed continuity of L at both s and t, and convergence of L_n to L at the continuity points, so by 4.8 and 4.9. ☐

Let us now have a look at the images of processes discussed in the previous lemma, so

$$Im\left(\tfrac{L_n(t)}{L_n(1)}\right)_{t\in[0,1]} \text{ and } Im\left(\tfrac{L(t)}{L(1)}\right)_{t\in[0,1]}.$$

These both are clearly subsets of the interval $[0,1]$, first represented by n dots, second – by line segments. So by "wrapping them around" and scaling (i.e. by applying $t \mapsto e^{2\pi i t}$ mapping) we indeed obtain subsets of unit circle: scaled version of cycle $\mathcal{C}_n$ (which we have already denoted by $\mathcal{G}_n$) and a random sum of arcs $\mathcal{G} = \left\{ e^{2\pi i \frac{L(t)}{L(1)}} : t \in [0,1] \right\}$.

The convergence we have just shown in Lemma 4.10 gives us therefore a fine candidate for the limit of $(\mathcal{G}_n, d_\mathcal{C})$, namely $(\mathcal{G}, d_\mathcal{C})$. However, we haven't yet computed Gromov-Hausdorff distance neither between mentioned images, nor between $\mathcal{G}_n$ and $\mathcal{G}$ (as these distances are not necessarily equal), so we shall still check that the known convergence of processes $\left(\frac{L_n(t)}{L_n(1)}\right)_{t\in[0,1]}$ implies the Gromov-Hausdorff convergence of spaces $(\mathcal{G}_n, d_\mathcal{C})$.

Let us present the proof of this fact as a simple combination of several lemmas. Of course, first of them is Lemma 4.10. As it claims that processes $\left(\frac{L_n(t)}{L_n(1)}\right)_{t\in[0,1]}$ and $\left(\frac{L(t)}{L(1)}\right)_{t\in[0,1]}$ are "close" to each other as the elements of $\mathcal{D}[0,1]$, it should also imply that already introduced images of these processes are "close" as the subsets of the interval $[0,1]$.

Lemma 4.12. For the random elements of $\mathcal{D}[0,1]$ such that $L_n \xrightarrow{a.s.} L$ we have

$$d_H\left(\left(Im\left(\frac{L_n(\cdot)}{L_n(1)}\right), d\right), \left(Im\left(\frac{L(\cdot)}{L(1)}\right), d\right)\right) \xrightarrow{a.s.} 0$$

where $d(x,y) = |x-y|$ denotes the usual metric on [0,1].

Proof. Let λ_n testify the convergence $L_n \xrightarrow{a.s.} L$ (provided by the Skorokhod's representation theorem). As, by definition, those λ_n are surjections, they do not affect the images:

$$Im\left(\frac{L_n(t)}{L_n(1)}\right)_{t\in[0,1]} = Im\left(\frac{L_n(\lambda_n(t))}{L_n(1)}\right)_{t\in[0,1]}$$

(thanks to that, we may actually think of λ_n only as of some disturbance of time when given value occurs).

Furthermore,

$$d_H\left(\left(Im\left(\frac{L_n(\lambda_n(\cdot))}{L_n(1)}\right), d\right), \left(Im\left(\frac{L(\cdot)}{L(1)}\right), d\right)\right) \leqslant \left\|\frac{L_n(\lambda_n(\cdot))}{L_n(1)} - \frac{L(\cdot)}{L(1)}\right\|_\infty$$

which is true directly from the definition of the Hausdorff distance: if we consider

$$r = \left\|\frac{L_n(\lambda_n(\cdot))}{L_n(1)} - \frac{L(\cdot)}{L(1)}\right\|_\infty,$$

then clearly

$$Im\left(\frac{L_n(\lambda_n(\cdot))}{L_n(1)}\right) \subseteq U_r\left(Im\left(\frac{L(\cdot)}{L(1)}\right)\right) \text{ and } Im\left(\frac{L(\cdot)}{L(1)}\right) \subseteq U_r\left(Im\left(\frac{L_n(\lambda_n(\cdot))}{L_n(1)}\right)\right)$$

and Hausdorff distance was defined as the infimum over all r fulfilling the above condition.

But, by the Lemma 4.10, we know that $\left\|\frac{L_n(\lambda_n(\cdot))}{L_n(1)} - \frac{L(\cdot)}{L(1)}\right\|_\infty \xrightarrow{n} 0$ which justifies the claim of this lemma. □

As the next lemma convinces, we actually already know enough. That means, if the two subspaces of the interval are close in the Hausdorff sense (which we already know), and we "wrap them around" and scale, they will remain close in the same sense – just as subspaces of the circle.

Lemma 4.13. Let $A, B \subseteq [0,1]$ be metric spaces with the usual metric on the interval. Then

$$d_H\left(e^{2\pi i A}, e^{2\pi i B}\right) \leqslant 2\pi d_H(A,B)$$

where $e^{2\pi i A}, e^{2\pi i B}$ are subspaces of the unit circle with the metric $d_{\mathcal{C}}$.

Proof. Let us take $r = d_H(A, B)$ i.e. the smallest satisfying $A \subseteq U_r(B)$ and $B \subseteq U_r(A)$. It is enough to show that

$$e^{2\pi iA} \subseteq U_{2\pi r}(e^{2\pi iB}) \text{ and } e^{2\pi iB} \subseteq U_{2\pi r}(e^{2\pi iA})$$

(in fact, by the symmetry, it is enough to show just one of the above inclusions which we shall now do).

Take $x \in e^{2\pi iA}$. Such x is of the form $e^{2\pi ia}$ for some $a \in A$. Therefore $d(a, b) < r$ for some $b \in B$ as $A \subseteq U_r(B)$. But that means $d_{\mathcal{C}}(e^{2\pi ia}, e^{2\pi ib}) < 2\pi r$ where, of course, $e^{2\pi ib} \in e^{2\pi iB}$. So $d_{\mathcal{C}}(x, e^{2\pi iB}) < 2\pi r$ and equivalently $x \in U_{2\pi r}(e^{2\pi iB})$.

□

We have been foreshadowing the main statement of this section a couple of times, so it might be already seen as a consequence of the last lemmas. However, for the completeness and clarity, let us now formulate it once again and combine the justified facts to argue for its truthfulness.

Theorem 4.14. Let $L_n(t) = n^{-1/\alpha} S_{[nt]}$ and denote its limit in $\mathcal{D}[0, 1]$ by L. Consider $X_n = \left(\mathcal{C}_n, \frac{2\pi d_{\mathcal{C}_n}}{S_n}\right)$ and $X = \left(\left\{e^{2\pi i \frac{L(t)}{L(1)}} : t \in [0, 1]\right\}, d_{\mathcal{C}}\right)$. For such metric spaces, as $n \to \infty$, we have

$$X_n \xrightarrow{\mathbb{P}} X$$

in the Gromov-Hausdorff sense.

Proof. Recalling the processes

$$\mathcal{G}_n = \left\{e^{2\pi i \frac{L_n(t)}{L_n(1)}} : \ t \in [0, 1]\right\} \text{ and } \mathcal{G} = Y,$$

by Claim 4.5 and Remark 2.3, for the proof of the theorem, it is enough to show the Hausdorff convergence $(\mathcal{G}_n, d_{\mathcal{C}}) \xrightarrow{\mathbb{P}} (\mathcal{G}, d_{\mathcal{C}})$, i.e.

$$d_H\left((\mathcal{G}_n, d_{\mathcal{C}}), (\mathcal{G}, d_{\mathcal{C}})\right) \xrightarrow{\mathbb{P}} 0.$$

The above is a consequence of the inequality provided by the Lemma 4.13:

$$d_H\left((\mathcal{G}_n, d_{\mathcal{C}}), (\mathcal{G}, d_{\mathcal{C}})\right) \leqslant d_H\left(\left(Im\left(\frac{L_n(t)}{L_n(1)}\right), d\right), \left(Im\left(\frac{L(t)}{L(1)}\right), d\right)\right)$$

which right hand side, by the Lemma 4.12, disappears when $n \to \infty$.

□

4.3 Deterministic normalization

Correspondingly to the finite mean case, again, we would like to replace stochastic normalization constant $\frac{2\pi}{S_n}$ in the metric considered on the graph by some deterministic constant. This matter seems to cause more trouble now as we no longer have the law of large numbers to approximate S_n with, so coming up with a deterministic substitute of the $\frac{2\pi}{S_n}$ is not that straightforward anymore.

Luckily, we have already studied S_n long enough to be able to foresee its behaviour also in the infinite mean case. In particular, we have Corollary 4.3 on our side and by its merit we may wish that $n^{1/\alpha}L(1)$ would be a good replacement. However, L is random itself, so it is no help for the stochastic normalization issue.

Therefore we shall stay with the $n^{1/\alpha}$ alone. Omitting the scaling by $L(1)$ should now clearly affect the proprieties of the limit object. So far we've succeeded in choosing normalization constants such that the circle visible in the limit was a unit one; it was possible because we took care to keep the girth of the cycle equal (in stochastic cases) or close (in deterministic case) to 2π. Presently, so considering the metric $\frac{2\pi d_{\mathcal{C}_n}}{n^{1/\alpha}}$ the girth becomes equal exactly $2\pi n^{-1/\alpha}S_n$, so by the 4.3 – close to $2\pi L(1)$. This shall be the circumference of the circle being the superset of the random sum of arcs which appears in the limit.

All of that is gathered and formalized in the

Theorem 4.15. Let, again, $L_n(t) = n^{-1/\alpha}S_{[nt]}$ and denote its limit in $\mathcal{D}[0,1]$ by L. Consider $X_n = \left(\mathcal{C}_n, \frac{2\pi d_{\mathcal{C}_n}}{n^{1/\alpha}}\right)$ and $X = \left(L(1)\left\{e^{2\pi i \frac{L(t)}{L(1)}} : t \in [0,1]\right\}, d_{\mathcal{C}}\right)$. For such metric spaces, as $n \to \infty$, we have

$$X_n \xrightarrow{\mathbb{P}} X$$

in the Gromov-Hausdorff sense.

The argument here is quite similar to the finite mean case: we again define a correspondence between the graph and the limit subset of the circle (let us from now on refer to this limit object as $L(1)\mathcal{G}$, in consistency with the previous notation). And again, we would like to associate this correspondence with a surjection and we want it to be "close to identity", so we divide $L(1)\mathcal{G}$ into smaller, random arcs such that every arc consists of these points on the circle which are the closest to the same vertex of the cycle (of course, closest with respect to $d_{\mathcal{C}}$). A function mapping elements of the same arc on the adequate vertex is now, clearly, a surjection and seems to be "as near to identity as possible" which suggests that the distortion of the associated correspondence should be small enough – which we shall formalize and compute.

Proof. Let us begin by taking λ_n which, as always, indicate the convergence of L_n to L in

$\mathcal{D}[0,1]$, and noting that, by similar argument to the one given in Claim 4.5,

$$\left(\mathcal{C}_n, \frac{2\pi d_{\mathcal{C}_n}}{n^{1/\alpha}}\right) \cong (L_n(1)\mathcal{G}_n, d_{\mathcal{C}})$$

for $L_n(1)\mathcal{G}_n = \left\{L_n(1)e^{2\pi i \frac{L_n(t)}{L_n(1)}} : t \in [0,1]\right\} = \left\{L_n(1)e^{2\pi i \frac{L_n(\lambda_n(t))}{L_n(1)}} : t \in [0,1]\right\}$ as those λ_n are surjections. Thanks to that we only need to define the correspondences $\mathcal{R}_n$ between $L_n(1)\mathcal{G}_n$ and $L(1)\mathcal{G}$ which we do via the condition:

$$(x,y) \in \mathcal{R}_n \iff \exists t \in [0,1]\left(y = L(1)e^{2\pi i \frac{L(t)}{L(1)}}, x = L_n(1)e^{2\pi i \frac{L_n(\lambda_n(t))}{L_n(1)}}\right).$$

The above formula means that a single point of $L_n(1)\mathcal{G}_n$ is corresponding to all of the points $L(1)e^{2\pi i \frac{L(t)}{L(1)}}$ for t for which $L_n(\lambda_n(t))$ is constant. It's worth noting here, that even though L_n is constant on the intervals of length $\frac{1}{n}$, namely for $t \in \left[\frac{k}{n}, \frac{k+1}{n}\right)$, the lengths of arcs produced by these constant values of L_n, so lengths of arcs from $L_n(1)e^{2\pi i \frac{L_n(\lambda_n(k/n))}{L_n(1)}}$ to $L_n(1)e^{2\pi i \frac{L_n(\lambda_n((k+1)/n))}{L_n(1)}}$, are given by the weights Y_k. Therefore the construction is indeed quite parallel to the the previous case where we've been dividing the full circle into arcs of random length.

Now, recalling the definition of distortion, we have:

$$\operatorname{dis}(\mathcal{R}_n) = \sup\{|d_{L_n(1)\mathcal{G}_n}(x,x') - d_{L(1)\mathcal{G}}(y,y')| : (x,y),(x',y') \in \mathcal{R}_n\} =$$
$$\sup_{t_1,t_2 \in [0,1]}\left|\left(d_{\mathcal{C}}\left(L_n(1)e^{2\pi i \frac{L_n(\lambda_n(t_1))}{L_n(1)}}, L_n(1)e^{2\pi i \frac{L_n(\lambda_n(t_2))}{L_n(1)}}\right)\right) - d_{\mathcal{C}}\left(L(1)e^{2\pi i \frac{L(t_1)}{L(1)}}, L(1)e^{2\pi i \frac{L(t_2)}{L(1)}}\right)\right|.$$

Considering $t_1 < t_2$ above, we get

$$\begin{aligned}
&d_{\mathcal{C}}\left(L_n(1)e^{2\pi i \frac{L_n(\lambda_n(t_1))}{L_n(1)}}, L_n(1)e^{2\pi i \frac{L_n(\lambda_n(t_2))}{L_n(1)}}\right) = \\
&\quad 2\pi L_n(1)\left(\left|\frac{L_n(\lambda_n(t_1))}{L_n(1)} - \frac{L_n(\lambda_n(t_2))}{L_n(1)}\right| \wedge \left|1 - \left|\frac{L_n(\lambda_n(t_1))}{L_n(1)} - \frac{L_n(\lambda_n(t_2))}{L_n(1)}\right|\right|\right) = \\
&\qquad 2\pi(L_n(\lambda_n(t_2)) - L_n(\lambda_n(t_1))) \wedge 2\pi(L_n(1) + L_n(\lambda_n(t_1)) - L_n(\lambda_n(t_2)))
\end{aligned}$$

and similarly

$$d_{\mathcal{C}}\left(L(1)e^{2\pi i \frac{L(t_1)}{L_n(1)}}, L(1)e^{2\pi i \frac{L(t_2)}{L(1)}}\right) = 2\pi(L(t_2) - L(t_1)) \wedge 2\pi(L(1) + L(t_1) - L(t_2)).$$

Therefore

$$\begin{aligned}
\operatorname{dis}(\mathcal{R}_n) = &2\pi \sup_{t_1<t_2}(|L_n(\lambda_n(t_2)) - L(t_2) + L_n(\lambda_n(t_1)) - L(t_1)| \vee \\
&|L_n(1) - L(1) + L_n(\lambda_n(t_1)) - L(t_1) + L_n(\lambda_n(t_2)) - L(t_2)|) \leqslant \\
&2\pi\left(2\sup_{t_1}|L_n(\lambda_n(t_1)) - L(t_1)| + 2\sup_{t_2}|L_n(\lambda_n(t_2)) - L(t_2)| + |L_n(1) - L(1)|\right) \xrightarrow{n} 0
\end{aligned}$$

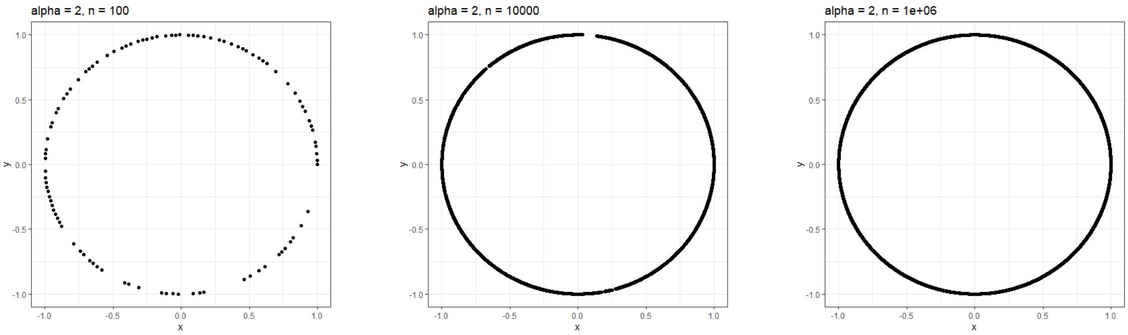
by the uniform convergence $L_n \circ \lambda_n \rightrightarrows L$ which follows from the definition of convergence in $\mathcal{D}[0,1]$.

□

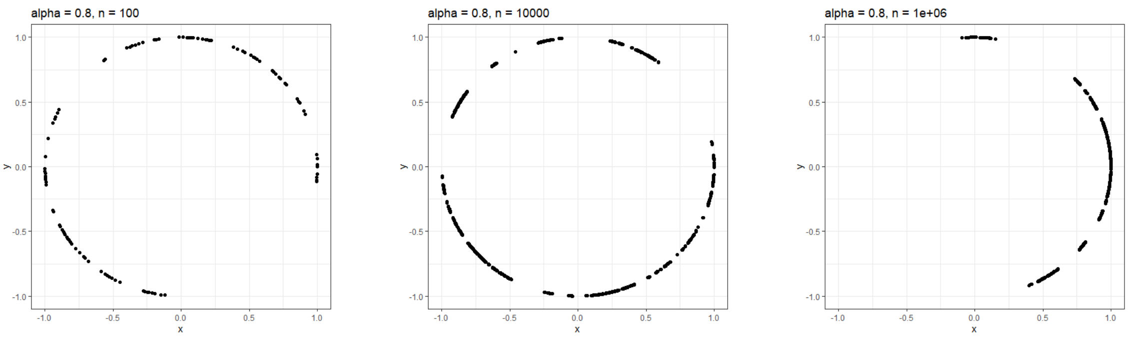
5 Visualisations and possible extensions

In a couple of various cases we have stated and proven that the cycle graph converges to the circle or its subset, and among all the technical details we are now in danger of loosing a natural, picturesque interpretation of such theorem. To avoid this danger and summarize our accomplishments, we shall visualise how the $\mathcal{C}_n$ outlines $\mathcal{C}$ or $\mathcal{G}$ better and better with the growth of n.

Below we can see the case of finite mean, precisely for $\alpha = 2$.

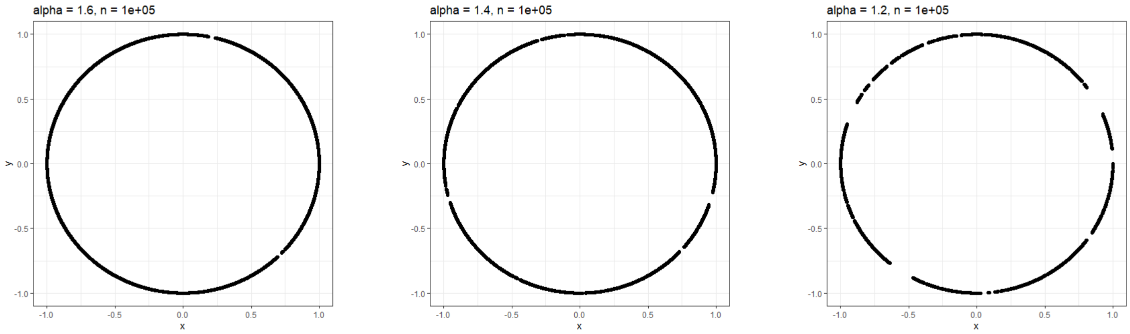


In the infinite mean case the limit object is the random sum of arcs. We simulate it for $\alpha = 0.8$.

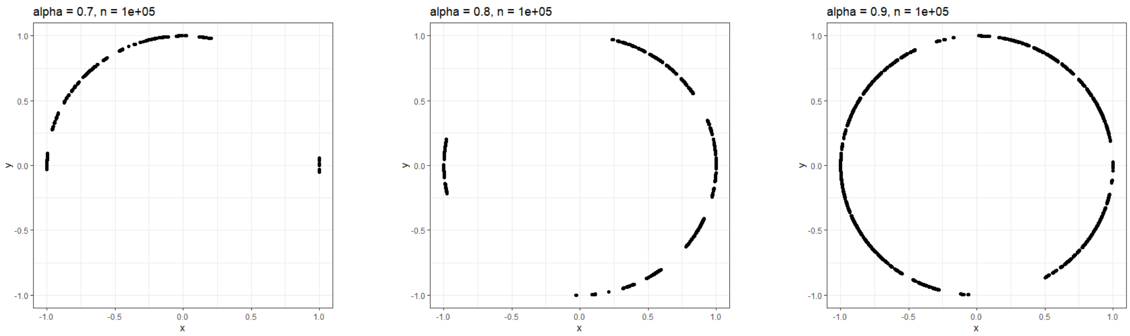


The observations confirm statements of the theorems which means that the outline of the limit objects becomes clearer when n grows larger. However, as the results for $n = 1000$ are not yet very satisfying, we assume that the rate of convergence is not too high.

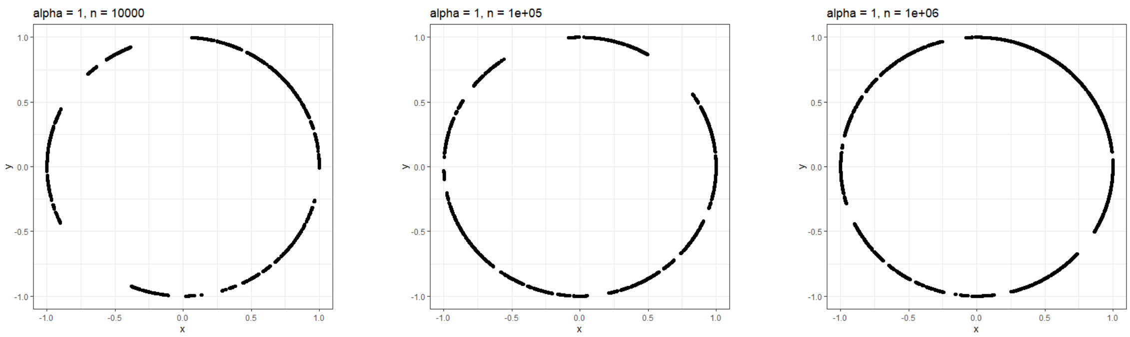
Another interesting question that might be asked having the plots to study, is the behaviour of the limit when $\alpha \to 1$ from above:



or from below:



Analyzing the plots, we might have reasons to assume that this limit with $\alpha \to 1$ should also be a unit circle. Still, we are not able to see it undoubtedly, as we can only simulate graphs with too few vertices:



References

[1] Billingsley, Patrick. *Convergence of probability measures.* John Wiley & Sons, 2013.

[2] Burago, Dmitri, Yuri Burago, and Sergei Ivanov. *A course in metric geometry.* Vol. 33. American Mathematical Society, 2022.

[3] Durrett, Rick. *Probability: theory and examples.* Vol. 49. Cambridge university press, 2019.

[4] van der Hofstad, Remco. "Stochastic processes on random graphs." *Lecture notes for the 47th Summer School in Probability Saint-Flour 2017* (2017).

[5] Resnick, Sidney I. *Heavy-tail phenomena: probabilistic and statistical modeling.* Springer Science & Business Media, 2007.

[6] Skorokhod, Anatolii V. "Limit theorems for stochastic processes." *Theory of Probability & Its Applications* 1.3 (1956): 261-290.