Google maps and improper Poisson line processes

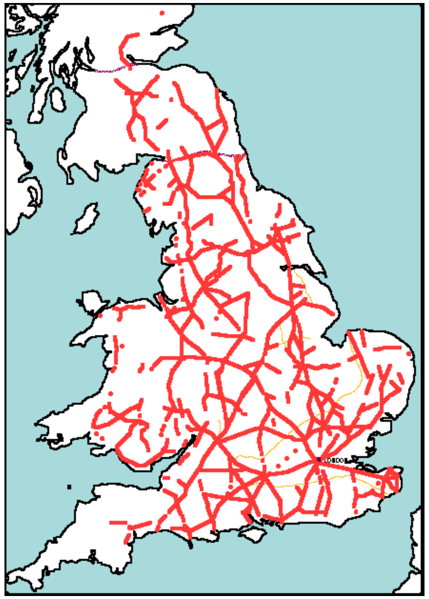
Simons Conference on Networks and Stochastic Geometry
University of Texas at Austin

Wilfrid Kendall (w.s.kendall@warwick.ac.uk)

18th May 2015

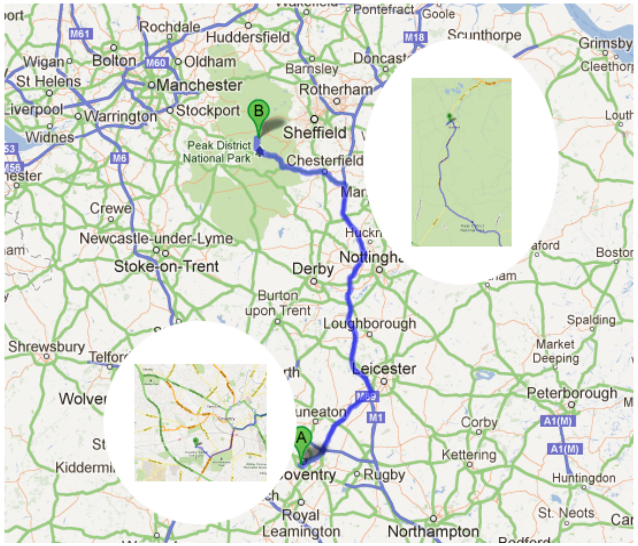
Frustrated optimization

for Roman Roads



Google holidays

for cats



Improving a network

for example, Roman Roads

- Network connecting N cities in rectangle side $\sqrt{N}$.

A Measure efficiency by minimizing connecting stuff? (Steiner tree)

B Measure efficiency by average excess of connection distance over Euclidean? (Complete planar graph)

- Aldous and WSK (2008): start with Steiner tree:
  - Add sparse set of random lines;
  - Add sparse rectilinear grid connecting lines and tree;
  - Add some box structures to avoid hotspots.
- Resulting network (large N) is economical with connection stuff, but the average excess is only logarithmic in N.
- Debunks a "natural" statistic for network efficiency.

(But see Aldous and Shun 2010.)

Further results:

1. Logarithmic upper bound for "mean connection distance minus Euclidean distance" using Poisson line process.
   (Steinhaus estimator for distance, study intersections of original Poisson line process and independent copy, focus on a certain Poisson polygon);
2. Proof that logarithmic upper bound is of correct order.
   (Stereological estimation using 50-year-old refinements of Mills ratio inequality);
3. Controlling fluctuations by bounding variance of excess connection distance.
   (represent perimeter of a Poisson polygon using theory of Lévy processes and self-similar processes).
4. Statistics of flows in the network based on the invariant Poisson line process.
   This uses an unusual anisotropic improper Poisson line process.

Random Line Processes

Brief description of geometry of Poisson line processes:

Poisson line processes in $\mathbb{R}^d$:

- Parametrize by ϖ "direction" of (undirected) line (point on "hemisphere" – actually, projective space!), and x location on perpendicular hyperplane.
- Invariant measure $c_d \, \mathrm{d}\, x \times \nu_{d-1}(\mathrm{d}\, \varpi)$.
  Coordinate x "twisted" by ϖ: unseen by measure theory.
- Variant parametrization replaces x by p, intersection of ℓ with reference hyperplane.
  Invariant measure now $c_d \sin\theta \; \mathrm{d}\, p \times \nu_{d-1}(\mathrm{d}\, \varpi)$.

Scale-invariant Random Spatial Networks (SIRSN)

Google holidays for cats

Aldous (2014) axiomatic approach:

- Input: set of nodes $x_1, \ldots, x_n$;
- Output: random network $N(x_1, \ldots, x_n)$ connecting nodes.

(1) Scale-invariance: $\mathcal{L}(N(\lambda x_1, \ldots, \lambda x_n)) = \mathcal{L}(\lambda N(x_1, \ldots, x_n))$ for each Euclidean similarity λ.

(2) Let D_1 be length of fastest route between two points at unit distance apart. We want $\mathbb{E}[D_1] < \infty$.

(3) Weak SIRSN property: consider network connecting points of an independent unit intensity Poisson point process. Average length per unit area of resulting "fastest route" network should be finite.

Models for SIRSN

SIRSN axioms have many interesting consequences. Models need to be hierarchical in some sense (fast versus slow). Paths exhibit "portal-like" behaviour.

Examples:

- Hierarchical binary model (randomized direction and location; Aldous 2014; Aldous and Ganesan 2013);
- Improper Poisson line process (also proposed by Aldous).

Improper Poisson line process model

Line process is marked by random speeds, and defined by a σ-finite measure:

- each line marked with positive speed-limit v;
- representing space is now parametrized by v, r, θ (more generally, in d dimensions, v, x, ϖ);
- to achieve scale-invariance, invariant measure is $\frac{\gamma-1}{2} v^{-\gamma} \, d\, v \, d\, r \, d\, \theta$ for $\gamma > 1$ (more generally, $c_d(\gamma-1)v^{-\gamma} \, d\, v \, d\, r \times \nu_{d-1}(d\, \varpi)$).
- The line process is dense throughout the plane (respectively, $\mathbb{R}^d$), but lines of speed exceeding threshold v_0 form proper Poisson line process if $\gamma > 1$ ($\gamma > d$).

  Use lines to go from A to B as fast as legally possible.

  For which γ might we get a decent network?

What is a path? (I)

Seek shortest-time paths ("temporal geodesics" or Π-geodesics) built using line process Π.
Require $\gamma > 1$, or fast lines will go everywhere.

- Introduce maximum speed limit, upper-semi-continuous $V : \mathbb{R}^d \to [0, \infty)$.
- A Π-path is locally Lipschitz, integrates measurable orientation field determined by Π, obeys speed limit.
- If $\gamma > d$ then:
  - there is an *a priori* random bound on distance travelled by Π-path in fixed time;
  - space of paths up to time T is closed, weakly closed in Sobolev space $L^{1,2}([0,T) \to \mathbb{R}^d)$,
  - paths up to time T, beginning in a compact set, together form a weakly compact set.

Consequence: Π-geodesics exist if Π-paths exist.

What is a path? (II)

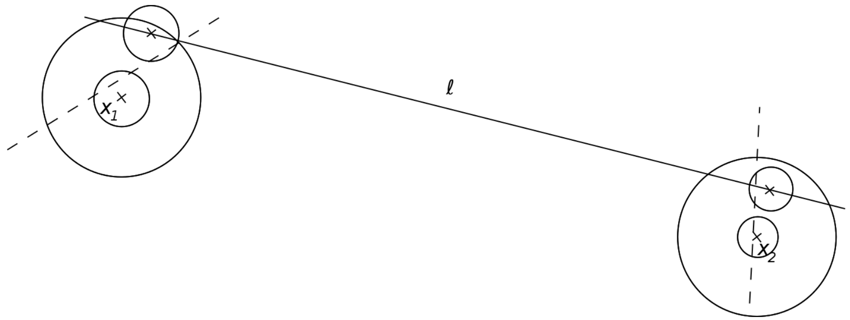
Suppose one wishes to connect two points ξ_1 and ξ_2 in $\mathbb{R}^d$ by a Π-path. Suppose $\gamma > d$ (turns out to be essential).

- Construct small balls around ξ_1 and ξ_2;
- Connect balls by fastest line ℓ intersecting both balls;
- Construct daughter nodes on ℓ closest to ξ_1 and ξ_2;
- Recurse.
- Borel-Cantelli, *et cetera*: establish almost sure existence of resulting path.
- This yields a binary tree representation of the path. Note that this is unavoidable if $d > 2$!

  A similar but more complicated argument almost surely allows simultaneous construction of paths between all possible pairs ξ_1 and ξ_2 in $\mathbb{R}^d$.

  Exercise: Visualize such paths in case $d = 3$.

Construction to connect two points



What is a path? (III)

Exponential moments for powers of Π-diameters

Kahn (2015), improving on WSK (2015, Theorem 3.6) has proved results of which the following is a simple example:

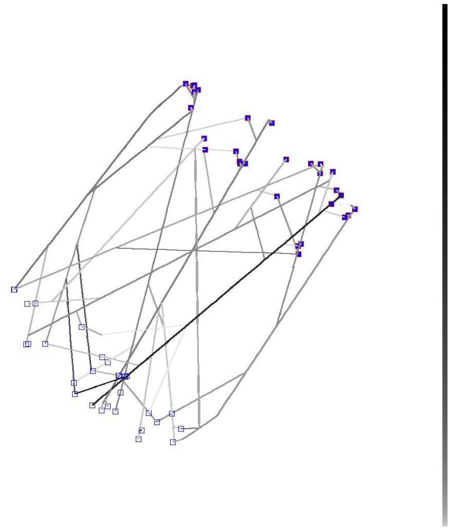
- Consider ball$(\mathbf{o}, R)$, a ball of Euclidean radius R.
- Let T_R be the supremum of the minimum times for a Π-path to pass from one point of ball$(\mathbf{o}, R)$ to another;

$$T_R = \sup_{x,y \in \text{ball}(\mathbf{o},R)} \inf\{T :$$
$$\text{some } \Pi\text{-path } \xi : [0,T] \to B \text{ satisfies } \xi(0) = x, \xi(T) = y\}.$$

  We call T_R the Π-diameter of ball$(\mathbf{o}, R)$.
- For explicit $\delta_R > 0$, $K_R < \infty$,

$$\mathbb{E}\left[\exp\left(\delta_R T_R^{\gamma-1}\right)\right] \quad \leq \quad K_R \quad < \quad \infty.$$

Simulations (approximate!) of a typical set of routes



Case $\gamma = 16$

Are Π-geodesics unique? (I)

Suppose now $d = 2$ and $\gamma > 2$, and we fix ξ_1 and $\xi_2 \in \mathbb{R}^2$. If Π is to generate a network between a finite set of points, then we need to know the Π-geodesic between ξ_1 and ξ_2 is almost surely unique.

- Theorem: All non-singleton intersections of Π-geodesic with lines ℓ of Π are “line meets line”.
  - First, reduce to case of ℓ being fastest line in region, with speed w.
  - Now change focus from high speed v to low “cost”, where

$$\text{“cost”} \quad = \quad \frac{\csc\theta}{v} - \frac{\cot\theta}{w}.$$

    where θ is angle of line with ℓ.
  - Argue that Π-geodesic hits ℓ using line of finite cost.

Are Π-geodesics unique? (II)

So Π-geodesics between ξ_1 and ξ_2 are made up of countable collection of intervals of lines of Π.

- Fix a given ℓ from Π, and consider the set S of such intervals lying in ℓ.
  - Consider two different finite collections $S_1 \subset S$ and $S_2 \subset S$, each composed of non-overlapping intervals.
  - Probability density argument: the total lengths of S_1 and S_2 have a joint density, unless one is empty.
  - Conditioning on time spent off ℓ, almost surely two Π-paths using S_1 and S_2 respectively must have different total travel times.
  - Almost surely two Π-geodesics between ξ_1 and ξ_2 must use the same finite collection of non-overlapping intervals from each ℓ of Π.
  - But we can reconstruct the Π-geodesic uniquely from the collections of intervals of each line ℓ in Π.
- Theorem: given ξ_1 and ξ_2, almost surely there is just one Π-geodesic between them.

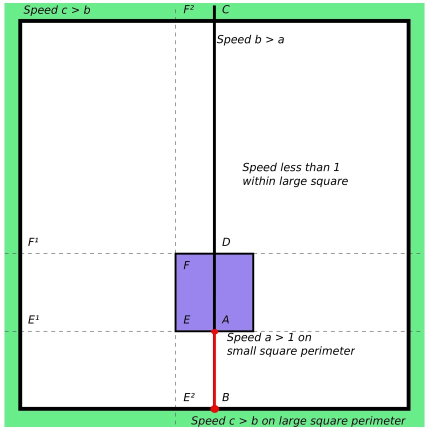
Do Π-geodesics have finite mean length?

Suppose again that $d = 2$ and $\gamma > 2$.

- Techniques for showing existence of Π-paths show finite mean of length of Π-geodesic *if lying in a fixed ball.*
- Could fast geodesics generate long lengths outside balls? (Oxford → Cambridge by motorway *via* London? or *Edinburgh*? . . .)
- Time spent by Π-geodesic can be bounded above by time spent on a circuit of a “racetrack” construction around ξ_1 and ξ_2 using fastest lines.
- We can upper-bound distance travelled outside a ball by using the “idealized path” construction employed above.
- The resulting perpetuity can be combined with the “racetrack” bound to establish finite mean length.

Diagram to prove Pre-SIRSN property

- A careful construction, together with a Borel-Cantelli argument, shows that re-scaled small perturbations of the following diagram (for suitable re-scaled speeds $1 < a < b < c$) can be found at all length scales:



If $c > 10b > 59a/3 > 354/3$, red segment is only exit.

Kahn (2015) has now produced a proof of the strong SIRSN property for all dimensions d and all $\gamma > d$, using the exponential inequality mentioned above for powers of Π-diameters

$$\mathbb{E}\left[\exp\left(\delta_R T_R^{\gamma-1}\right)\right] \quad \leq \quad K_R \quad < \quad \infty \, .$$

This can be used to show that Π-geodesics must make substantial re-use of shared lines: for all Π-geodesics bridging a suitable annulus, the exponential inequality forces each Π-geodesic to make substantial use of a limited number of “fast” lines. SIRSN follows by using a measure-theoretic version of the pigeonhole principle.
Kahn (2015) deploys further ingenious arguments to obtain uniqueness (essential if above argument is to work) and finite mean-length of Π-geodesics in case $d > 2$.

Conclusion

- We obtain scale-invariant random metric spaces in $\mathbb{R}^d$ for $\gamma > d$ (but visualizing paths in $d > 2$ is . . . interesting).
- The improper line process construction delivers a SIRSN for finite sets of points in the plane with $\gamma > 2$.
- Thanks to Jonas Kahn (2015), we even have the following Result: the SIRSN property holds in all dimensions (for $\gamma > d$), particularly *Average specific length of resulting “fastest long-distance route” network is finite.*
- Consider a Π-geodesic when $d = 2$. Can it ever come to a complete stop strictly between source and destination?
- Links to Brownian maps? Kahn (2015) highlights the Poisson line SIRSN with $\gamma = 3$, $d = 2$.

  Some reading: Aldous and WSK (2008), Aldous and Ganesan (2013), Aldous (2014), WSK (2011, 2014, 2015), Kahn (2015).

Aldous, D. J. (2014, April).
Scale-Invariant Random Spatial Networks.
Electronic Journal of Probability 19, no. 15, 1–41.

Aldous, D. J. and K. Ganesan (2013, December).
True scale-invariant random spatial networks.
Proceedings of the National Academy of Sciences of the United States of America 110(22), 8782–8785.

Aldous, D. J. and J. Shun (2010).
Connected spatial networks over random points and a route-length statistic.
Statistical Science 25(3), 1–28.

Aldous, D. J. and WSK (2008, March).
Short-length routes in low-cost networks via Poisson line patterns.
Advances in Applied Probability 40(1), 1–21.

Kahn, J. (2015).
Improper Poisson line process as SIRSN in any dimension.
arXiv 1503.03976, 34pp.

WSK (2008).
Networks and Poisson line patterns: fluctuation asymptotics.
Oberwolfach Reports 5(4), 2670–2672.

WSK (2011, October).
Geodesics and flows in a Poissonian city.
Annals of Applied Probability 21(3), 801–842.

WSK (2014).
Return to the Poissonian City.
Journal of Applied Probability 51A, 297–309.

WSK (2015, March).
From Random Lines to Metric Spaces.
Annals of Probability (to appear), 46.