Poissonian rain coloring and a self-similar process of coalescing planar partitions

David Aldous

February 22, 2017

Could be a 5-minute talk

- Here is a process that looks interesting.
- It exists.
- I would like to know more about it.

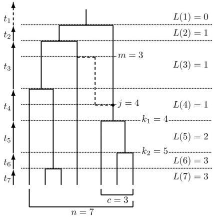
Having more time I will first give background, which perhaps fits the theme

stochastic coalescence meets stochastic geometry.

- Rapid overview of stochastic coalescence – stuff we knew ≥ 25 years ago.
- The natural geometric analogs – general problems I couldn't do 10 years ago and still can't do (and no-one else has thought about).
- The topic of this talk – a special case I can do. Proofs in *arXiv* paper.

3 conceptually different starting points for what I will call **coalescence** ("stuff joining up") and its opposite **fragmentation** ("stuff splitting up").

1. Most familiar within Applied Probability is the topic centered on the *Kingman coalescent* model; the objects are "lines of descent" traced backwards in real-world time.



2. An older setting envisages chunks of physical matter splitting or joining (e.g. polymers or colloids studied in physical chemistry). This has a huge literature going back to 1917.

Versuch einer mathematischen Theorie der Koagulationskinetik kolloider Lösungen.

Von

M. v. Smoluchowski.

(Mit 3 Figuren im Text.)

(Eingegangen am 8. 9. 16.)

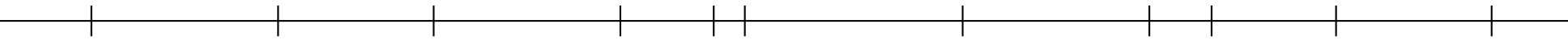
I. Einleitung.

So sehr auch bis heute die Literatur über Koagulation kolloider Lösungen angewachsen ist, sind doch unsere Kenntnisse betreffs des quantitativen Verlaufs, sowie betreffs des Mechanismus des Koagulationsprozesses äusserst mangelhaft. Die meisten Forscher begnügen sich mit qualitativen Beobachtungen oder stellen ihre Messungsreihen in Tabellen oder Kurvenform[1]) dar, da die mathematische Wiedergabe derselben auf aussergewöhnliche Schwierigkeiten stösst.

In den interessanten Arbeiten[2]) von S. Miyazawa, N. Ishizaka, H. Freundlich, J. A. Gann wird allerdings eine formelmässige Zusammenfassung des empirischen Versuchsmaterials, sowie eine Aufklärung desselben nach Analogie mit den Gesetzen der chemischen Kinetik an-

3. As an intermediate viewpoint, consider random splitting or joining of mathematical objects. Here is the simplest example.

Consider the infinite line as "space".
At time $-\infty < t < \infty$ there is a Poisson point process, rate e^t per unit length, of cuts. This defines intervals of Exponential(e^t) lengths.



Reinterpret lengths of intervals as masses of clusters in some unspecified (not d-dimensional) space. [board] As t varies what happens to these clusters?

As t **decreases**, the process fits the verbal formulation of the general (continuous-mass) *Smoluchowski coagulation equation*

> *a given cluster of mass x merges with some cluster of size y at rate $K(x, y)f_t(y)dy$ where $f_t(y)$ is the pdf of cluster masses at time t*

in the special case $K(x, y) = 2$.

Re-parameterize via $s = e^t$: at time $0 < s < \infty$ there is a Poisson point process, rate s per unit length, of cuts.

As s **increases**, the process fits the verbal formulation of the general (continuous-mass) *fragmentation process*

> *A cluster of mass x splits at rate x^α into two clusters of sizes Ux and $(1-U)x$*

in the special case $\alpha = 1$ and $U =_d$ Uniform$(0, 1)$.

Starting from this one example there are many directions one can go.

A. The mathematically simplest context for general models is fragmentation. In the classical scientific literature one writes down the collection of differential equations for

$$f(x, t) = \text{relative numbers of mass-}x\text{ clusters at time } t.$$

A nicer "probabilistic" approach (e.g. Brennan - Durrett 1987) is to observe that the the mass of the cluster containing a typical atom is an easily-described Markov process on $\mathbb{R}^+$. In particular, provided smaller clusters split more slowly than larger ones, under mild assumptions we get an asymptotic self-similarity result:

$$\text{time-}t\text{ size distribution of clusters } \sim t^{-\beta} X.$$

On the other hand, if smaller clusters split more quickly then (under mild assumptions) then the matter may be "reduced to dust" in finite time.

Studying **coalescence** with general kernels $K(x, y)$ is harder – the process "mass of the cluster containing a typical atom" is not a simply-described Markov process. But classical work implies a similar dichotomy:

if larger clusters merge more quickly, then infinite clusters appear in finite time – **gelation**

if not then we expect asymptotic self-similarity

$$\text{time-}t \text{ size distribution of clusters } \sim t^{\beta} X.$$

Note: 4 flavors of coalescing models.
Stochastic models with finite total mass; either continuous (total mass 1) or discrete (total mass N).
Deterministic models (informally, limits of expectations in the stochastic models) for densities in infinite volume:
either discrete: $f(i, t) =$ density of mass-i clusters
or continuous: $f(x, t)dx =$ density of mass-dx clusters.

B. In any model there is qualitative time-reversal duality; if one direction of time is a model of (binary) fragmentation then the other direction is a model of (binary) coalescence. But when is there an exact duality between simple models, as in our previous example?

A non-obvious example is provided by what we call **the additive Marcus-Lushnikov process and the additive coalescent.** The following construction is explicit in Pitman (1998), implicit in Yao (1976).

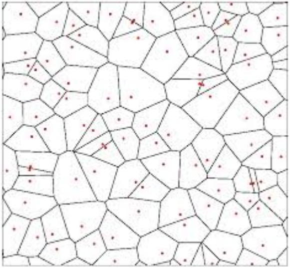
Take a random N-vertex tree, that is a uniform pick from the N^{N-2} trees on N distinguished vertices. Attach independent exponential(1) random variables ξ_e to each edge e. At each $0 \leq t < \infty$ there is a forest consisting of the edges $\{e : \xi_e \leq t\}$.

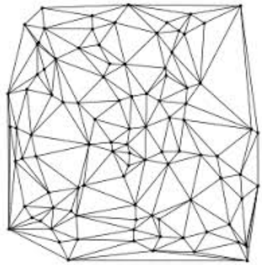
- At $t = 0$ there are N isolated vertices.
- At $t = \infty$ there is one component of size N.

Over $0 < t < \infty$ the process "sizes of connected components" evolves as the **additive** Marcus-Lushnikov process, that is the discrete stochastic coalescent with $K(x, y) = x + y$.

Part 2: are there interesting geometric (2-dimensions) analogs of coalescence/fragmentation?

Stochastic geometry contains many models of random partitions of the plane, typically based on a Poisson point process.





But the usual models do not extend directly to mathematically natural models of binary coalescence/fragmentation, because adding another point affects nearby regions. As a theory project, can we find tractable models of coalescence/fragmentation of partitions in the plane?

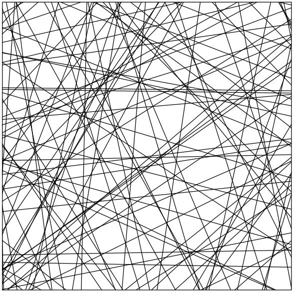
My point in giving the classical non-geometric theory background is that it suggests

1. Models of fragmentation should be easier to study.
2. Look for special processes with exact coalesce/fragment duality.
3. Coalescence models should show dichotomy between "percolation" (infinite clusters appear in finite time) and "self-similarity" (limit distribution exists under rescaling).

Where does this get us ???

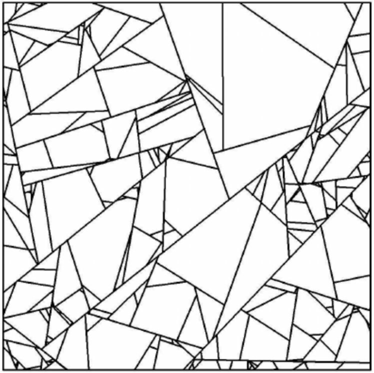
- Models of fragmentation should be easier to study.

One could use the Poisson line process



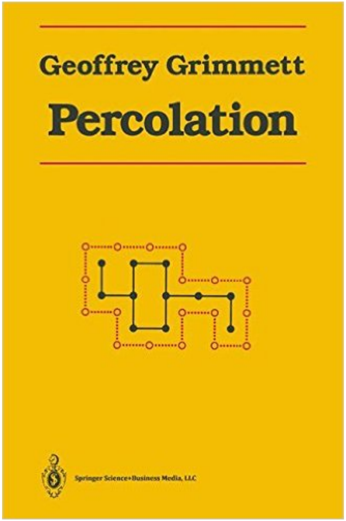
but this is “non-local” and not binary.

There is a class of STIT models in which regions are independently split via some scale-invariant rule. As in the non-geometric setting one can study "the region containing a typical point" as an autonomous Markov process.

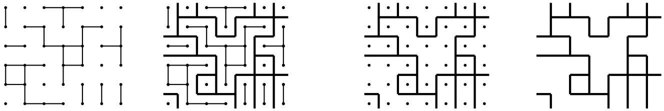


However the time-reversed coalescence process is essentially deterministic.

- Can we find special processes with exact stochastic coalesce/fragment duality?


Geoffrey Grimmett
Percolation
Springer Science+Business Media, LLC

Any set of edges in the lattice $\mathbb{Z}^2$ determines a partition of $\mathbb{R}^2$ via planar duality. So take the classical bond percolation process – edges appear at Exponential(1) times – and consider it as a partition-valued process.



It evolves according to the general rule

adjacent regions A, B merge at rate $K(A, B)$

in the special case

$$K(A, B) = \text{ length of boundary between } A \text{ and } B.$$

So 10 years ago I thought about the general class of models

adjacent regions A, B merge at specified rate $K(A, B)$

Intuitively we expect the same dichotomy:
• if larger regions merge sufficiently faster than smaller ones then infinite regions appear in finite time – **percolation**
• if not then asymptotic self-similarity.

But that project was ....$\mathbb{EPIC}$ $\mathbb{FAIL}$.........couldn't prove anything non-trivial.

• for $K \equiv 1$ intuitively clear there is percolation (Peierls contours *almost* works)
• couldn't prove any such process has a self-similar limit.

Simulations and discussion in *Empires and percolation: stochastic merging of adjacent regions* (2010).

Part 3: On to the new material

Instead of visualizing countries merging into empires, let us visualize countries with capital cities. Here is a model. Specify that

- at time t the capital cities form a rate e^t PPP

with the natural coupling: as t decreases cities disappear at stochastic rate 1. What happens to the countries? We specify

- when a city is deleted, its country is appended to the country

whose capital city is closest to the deleted city.

Because the PPP of cities is scale-invariant, it is intuitively very plausible that the partition-valued process has asymptotic self-similarity, implying

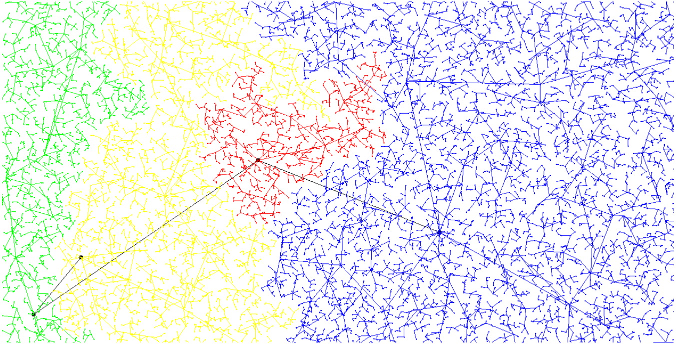
> *there exists a self-similar version of the partition-valued process over time $\infty > t > -\infty$.*

This is **Theorem 2** in the paper. But we don't know much about this process, quantitatively or qualitatively.

This model arose from a (superficially) different model, as follows.

Choose $k \geq 2$ distinct points $z_1, \ldots, z_k$ in the unit square, and assign to point z_i the color i from a palette of k colors. Take i.i.d. uniform random points $U_{k+1}, U_{k+2}, \ldots$ in the unit square, and inductively, for $j \geq k+1$,

> *give point U_j the color of the closest point to U_j amongst $U_1, \ldots, U_{j-1}$ where we interpret $U_i = z_i, 1 \leq i \leq k$.*



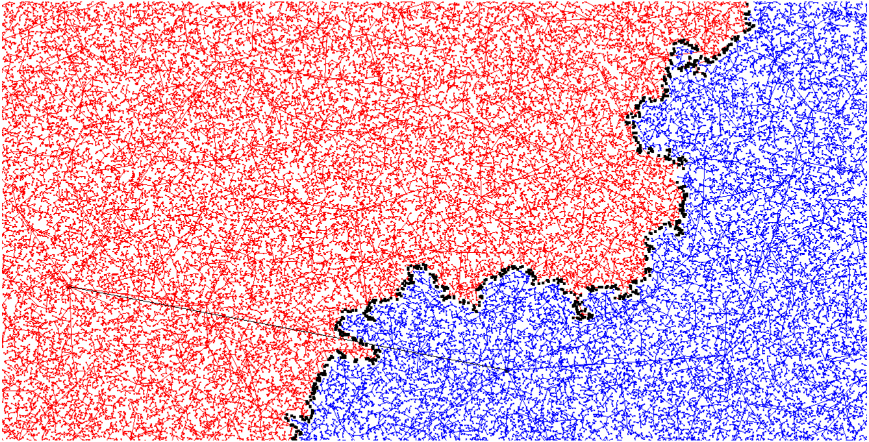
Simulations and intuition strongly suggest that there is (in some sense) an $n \to \infty$ limit which is a random partition of the square into k colored regions. This "coloring model" was considered independently by several people over last 10 years, but no results.

At first sight the convergence assertion seems easy.

First consider Voronoi regions. Intuitively, the area of the Voronoi region of a given color should behave almost as a martingale, because a new particle near the boundary seems equally likely to make the area larger or smaller. If one could bound the martingale approximation well enough to establish a.s. convergence of such areas, the convergence theorem would follow rather trivially. But doing so seems to require detailed knowledge of the geometry of the boundary.

We will give a different formalization (later) of this "convergence theorem" as **Theorem 1** of the paper.

Simulations also suggest that the boundaries between these limit regions should be fractal, in some sense.



It is technically convenient to work with a slightly more sophisticated version of the "Poisson coloring" model.

Consider the space-time "Poisson rain" process on $\mathbb{R}^2$ where the particles arriving before time t form a PPP of rate e^t per unit area. Consider a time t_1; assign different colors to the different particles present at time t_1. Then run the coloring process over $t_1 < t < \infty$.

Theorem 1: the convergence theorem says we can take the $t \to \infty$ limit to define a partition $\mathcal{A}(t_1)$ of $\mathbb{R}^2$ into countries with capital cities.

(outline later)

Consider the space-time "Poisson rain" process on $\mathbb{R}^2$ where the particles arriving before time t form a PPP of rate e^t per unit area. Consider a time t_1; assign different colors to the different particles present at time t_1. Then run the coloring process over $t_1 < t < \infty$.
Theorem 1: the convergence theorem says we can take the $t \to \infty$ limit to define a partition $\mathcal{A}(t_1)$ of $\mathbb{R}^2$ into countries with capital cities.

What is the relation between $\mathcal{A}(t_1)$ and $\mathcal{A}(t_1 - dt)$????

If a new particle arrives at z during $[t_1 - dt, t_1]$ with nearest previous particle z', then the limit regions split as $A_{z'}(t_1 - dt) = A_{z'}(t_1) \cup A_z(t_1)$.

So in reversed time the regions merge as claimed.

- when a city is deleted, its country is appended to the country whose capital city is closest to the deleted city.

And self-similarity of this "coalescing partitions" process follows from self-similarity of the space-time PPP.

How to formulate the convergence theorem

Instead of using Voronoi regions we use a weaker formulation. At time t the particles form a PPP of rate e^t, so the (weighted) empirical measure that puts weight e^{-t} on each particle is a random measure that converges a.s. to Lebesgue measure as $t \to \infty$. So for each color c (representing a particle at a given time) there is the corresponding random measure $\mu_{c,t} = \mu_{c,t}(\omega, A),\ A \subset \mathbb{R}^2$ counting only the color-c particles. We prove that for each color c

- $\mu_{c,t} \to$ some limit random measure $\mu_{c,\infty}$ as $t \to \infty$ (in probability; weakly).
- $\mu_{c,\infty}$ is Lebesgue measure restricted to some random measurable set.

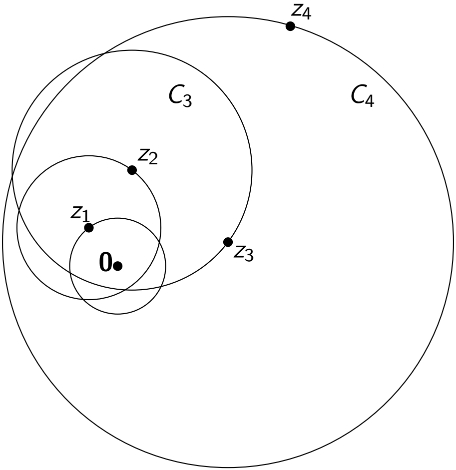
So we can define $\mathcal{A}(t_1)$ as the partition into measurable sets obtained in this way from coloring the particles present at time t_1.

Note this does not give topological information about the regions.

Outline of proof the convergence theorem

- Genealogical tree structure: arriving particle is "child" of nearest previous particle.
- w.l.o.g take $t_1 = 0$; need to show that particles which are very close at a large time t have w.h.p. the same color, that is same time-0 ancestor.
- By self-similarity, this is equivalent to showing that for particles at $O(1)$ distance at time 0, their "lines of descent" merge at some time $-T$ and some distance R that we can bound.

So we need to study lines of descent.


z_4
C_3
C_4
z_2
z_1
z_3
$\mathbf{0}$

The first part of the proof shows that for a typical particle ξ present at time 0,

• the distance to the ancestor at time $-t$ is of order $e^{t/2}$,

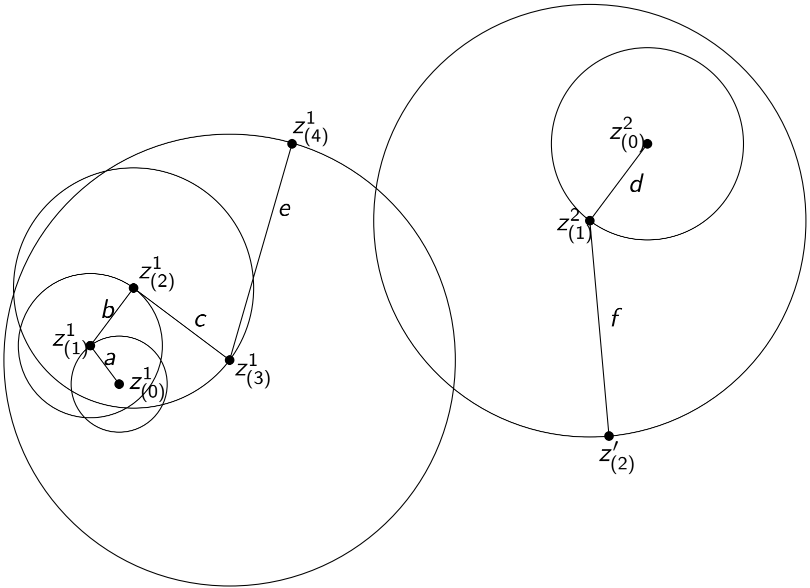
which is the same order as the distance to the *nearest* particle present at time $-t$.

In the second part of the proof we first consider two particles present at time 0 and distance r apart. Their lines of descent merge at some random past time $-T_{\mathrm{coal}}$, and we need an upper bound (next slide) on the tail of this distribution. The methods in these sections are very concrete – calculations and bounds involving Euclidean geometry and Poisson processes – though rather intricate in detail.

Proposition

There exist constants $K, \beta < \infty$ and $\rho > 0$ such that, in the "coupled lines of descent" process, for any r

$$\mathbb{P}(T_{\mathrm{coal}} > t) \leq K \exp(-\rho t) \text{ for all } t > \beta \log^+ r. \tag{1}$$


$z^1_{(4)}$
$z^2_{(0)}$
d
e
$z^2_{(1)}$
$z^1_{(2)}$
b
c
f
$z^1_{(1)}$
a
$z^1_{(3)}$
$z^1_{(0)}$
$z'_{(2)}$

The other ingredient is a coupling argument. Recall $\mu_{c,t}$ is the random measure putting weight e^{-t} on each time-t descendant of a given time-0 particle c. We need to show that for $0 \ll t_1 \ll t_2$

μ_{c,t_1} and μ_{c,t_2} are close (in probability), with distance $\to 0$ as $t_1 \to \infty$.

The argument is rather subtle because (analogy with supercritical branching processes) the distribution of time-t_1 ancestor of uniform time-t_2 descendant is not uniform. But once we know the limit random measures exist, Proposition 1 is enough to show that a limit random measure is Lebesgue measure on some random set.