Overview: known results, open problems

(A) There are 3 **special kernels** $K(x,y)$

1	$x+y$	xy
constant	additive	multiplicative

which are well understood. Solutions of the Smoluchowski coagulation equation have been obtained analytically; explicit constructions of the stochastic coalescent have been made. Most questions about these cases can mow be answered by calculation.

(B) **Plausible heuristics** for general $K(x,y)$, based on self-consistency:
(i) self-similar scaling laws
(ii) conditions for gelation phase transition
(iii) time for giant clusters to appear in the Marcus-Lushnikov process.

(C) **Recent questions** posed by mathematicians.

(i) Precise interpretation of gelation within stochastic models
(ii) Entrance boundaries for the stochastic coalescent.

(D) **Foundational** math questions

(i) Existence and uniqueness of solutions of Smoluchowski coagulation equation
(ii) WLLN connecting Marcus-Lushnikov to Smoluchowski coagulation equation
(iii) regularity properties of the stochastic coalescent (Feller process?)

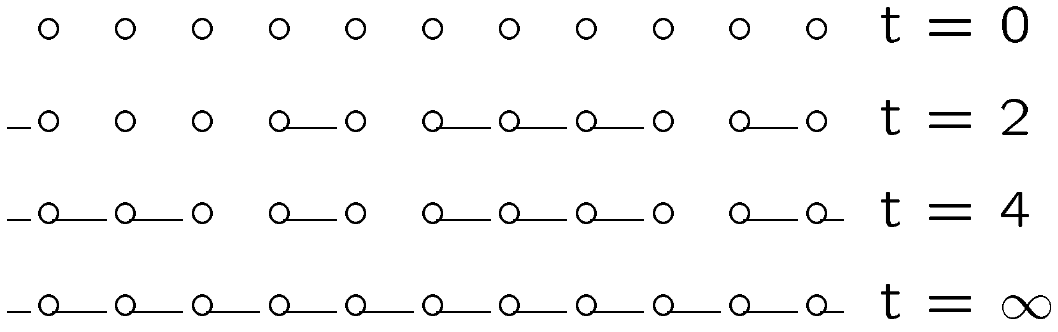
Given some special continuous-time Markov process, it has distributions at times t, and these satisfy some differential equations. Look at this the other way round; the solutions of certain differential equations can be interpreted as the marginal distributions of some special process. Prototype instance

- Brownian motion solves the heat equation.

One can invent such "process solutions" of the Smoluchowski coagulation equation with each of the 3 special kernels. We'll just discuss the $K(x, y) = 1$ case.

Recall interpretation of SCE:
a tagged mass-x cluster merges with some mass-y cluster at rate $K(x, y)n(y, t)$.

Construction of a stochastic process solving $K(x,y) = 1$ Smoluchowski



Start with unit mass at each integer site. Edges appear at independent times T, $P(T > t) = G(t)$. Average number of x-clusters per unit length at time t is

$$n(x,t) = G^2(t)\ (1 - G(t))^{x-1}.$$

A tagged cluster of mass x will merge with a neighboring cluster of some mass y with rate

$$2\ \frac{-G'(t)}{G(t)}\ \frac{n(y,t)}{G(t)}$$

Choosing $G(t) = 2/(2+t)$ makes this rate $1 \cdot n(y,t)$. So this $n(x,t)$ solves the discrete Smoluchowski equations for $K(x,y) = 1$.

In the Smoluchowski coagulation equation we guess that as $t \to \infty$ the distribution of cluster sizes, relative to its mean $s(t)$, has some limit density $\psi(x)$; precisely

$$n(x,t) \approx \frac{1}{s^2(t)}\ \psi\left(\frac{x}{s(t)}\right) \qquad (*).$$

If the kernel K is self-similar with exponent $\gamma < 1$ then we have self-consistency with $s(t) \approx t^{1/(1-\gamma)}$ and with $\psi(\cdot)$ satisying

$$\frac{1}{2}\int_0^x K(y,x-y)\psi(y)\psi(x-y)dy$$

$$-\psi(x)\int_0^\infty K(x,y)\psi(y)dy$$

$$= w\{2\psi(x) + x\psi'(x)\} \text{ ; some } w > 0.$$

This calculation goes back to Friedlander and Wang (1966). But no rigorous proof of:

- given K, does ψ exist? unique?
- does convergence (*) hold?

In this $K(x, y) = 1$ case

$$n(x,t) \sim \frac{4}{t^2} \exp(-\frac{2x}{t}).$$

Fits the general scheme (next slide) with

$$s(t) = t/2; \quad \psi(x) = \exp(-x).$$

Comments.

1. For scaling exponent $\gamma = 1$ same heuristics say typical cluster size $s(t) \approx e^{ct}$.

2. In the example of coalescence for particles undergoing physical Brownian motion, $K(x, y)$ has scaling exponent $\gamma = 0$. Theory predicts mean mass of clusters is $s(t) \approx t^1$ and so the typical radius is $r(t) \approx t^{1/3}$.

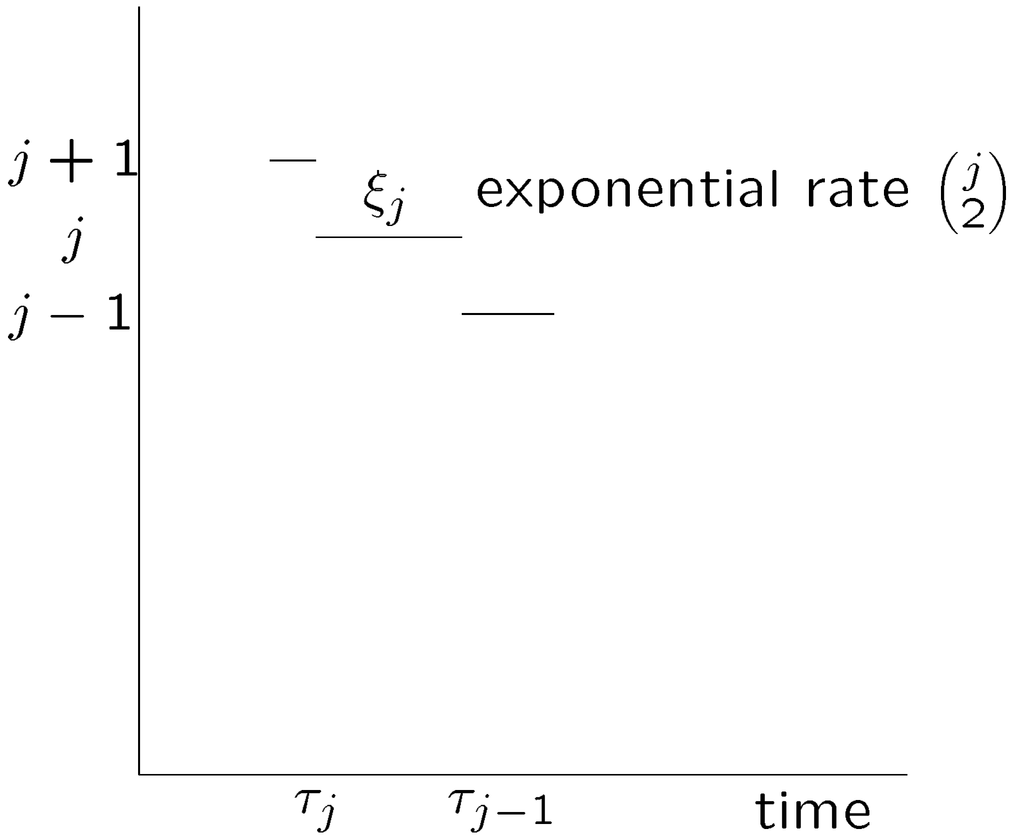
3. Consequences for Marcus-Lushnikov are that time until some "giant" (size of order N) cluster appears should be

$$\approx N^{1-\gamma} \quad (\gamma < 1)$$

$$\approx \log N \quad (\gamma = 1).$$

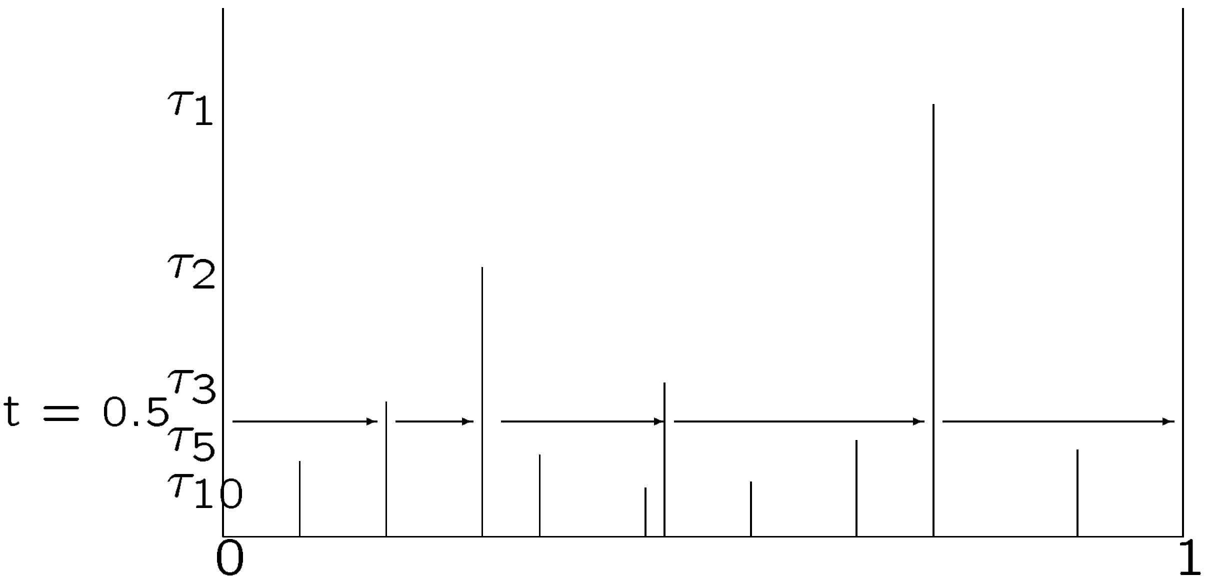
In trying to make a rigorous proof, difficulty is to show that relative size distribution stays tight as t increases; proved only in some $\gamma < 0$ cases by Clark - Katsouros (1999).

The $K(x,y) = 1$ stochastic coalescent is essentially Kingman's coalescent from mathematical population genetics. It has the special property that
(number of clusters at time t) is Markov.



$$\tau_j = \sum_{i=j+1}^{\infty} \xi_i < \infty.$$

To picture the $K(x,y) = 1$ stochastic coalescent, identify mass of clusters $\leftrightarrow$ lengths of subintervals of $[0,1]$.



In reversed time, the positions of split-points are independent, uniform on $[0,1]$.
This constructs a version starting with infinitesimally small clusters at time 0; not hard to show uniqueness.

Note different time scalings: in $K(x,y) = 1$ case

M-L and SCE	Stochastic coalescent
t	t/N

For $t \to \infty$, $t/N \to 0$ we can check consistency. SCE predicts typical cluster size is exponential, mean $\sim t/2$. Stochastic coalescent predicts typical cluster size is $N \times$ exponential, mean $1/m(t)$, where $m(t)$ is mean number of clusters in Kingman's coalescent. Since

$$E\tau_j = \sum_{i=j+1}^{\infty} 1/\binom{i}{2} \sim 2/j$$

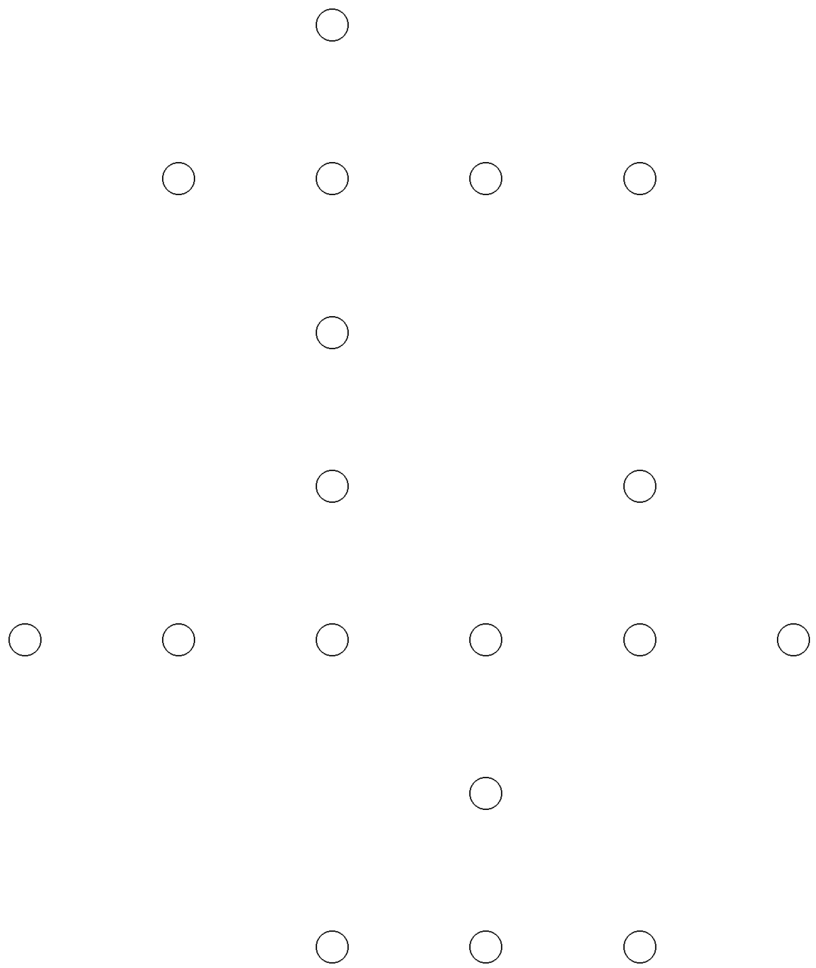
we find $m(t) \sim 2/t$ as $t \downarrow 0$. So typical cluster size predicted by stochastic coalescent is $tN/2$.

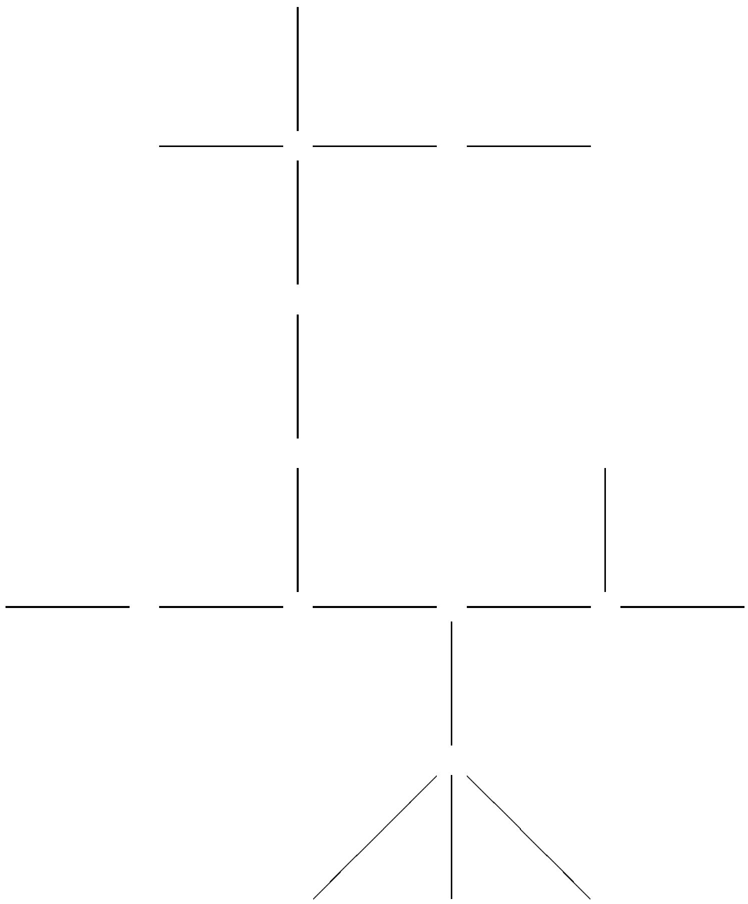
The additive Marcus-Lushnikov process and the additive coalescent.

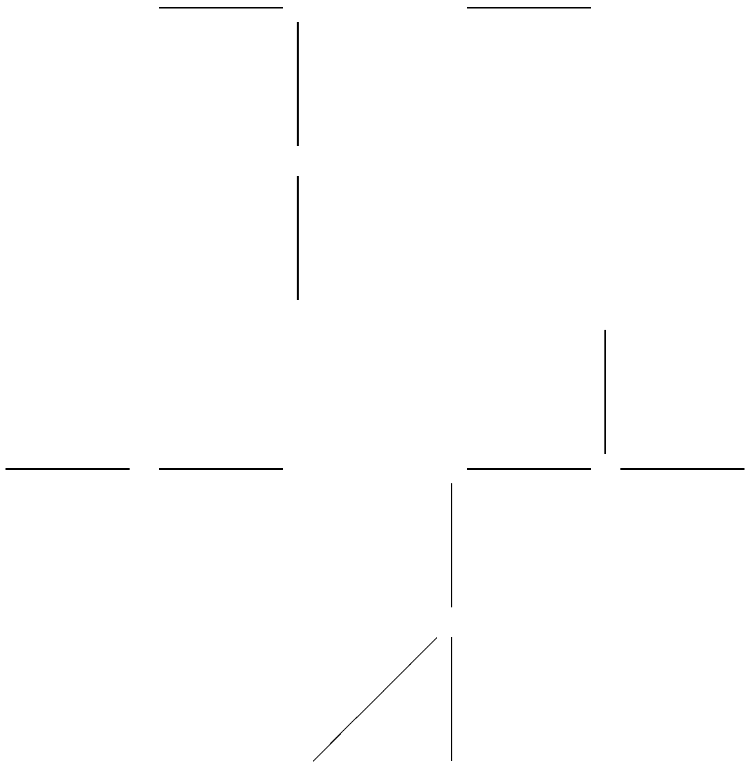
Construction. (explicit in Pitman (1998), implicit in Yao (1976)). Take a random N-vertex tree, that is a uniform pick from the N^{N-2} trees on N distinguished vertices. Attach independent exponential(1) random variables ξ_e to each edge e. At each $0 \leq t < \infty$ there is a forest consisting of $\{e : \xi_e \leq t\}$.

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 $K(x, y) = x + y$

Remarkably, one can let $N \to \infty$ in this construction.







In the random N-vertex tree, the distance between two random vertices is order $N^{1/2}$. Rescale edge-lengths to have length $N^{-1/2}$; let $N \to \infty$ to obtain the continuum random tree (CRT).

The figure shows a simulation of the subtree of the CRT spanned by 50 random vertices. A realization of the CRT has a mass measure μ of total mass 1; the 50 leaves in the figure are i.i.d. samples from the mass measure. The realization also has a σ-finite length measure $\ell(\cdot)$ on the skeleton.

Can construct the CRT by a deterministic map from Brownian excursion, but that's not so helpful for our purposes.

For each $0 \leq \lambda < \infty$, take a Poisson process of **cuts** $*$ along the skeleton, rate λ per unit $\ell(\cdot)$-length. These cuts partition the tree into components; write

$$\mathbf{Y}(\lambda) = (Y_1(\lambda), Y_2(\lambda), \ldots)$$

for the μ-masses of the components.

Aldous - Pitman (1998) show that the process $(\mathbf{X}(t) = \mathbf{Y}(e^{-t}), -\infty < t < \infty)$ evolves as the **additive coalescent**. The mass $X_1(t)$ of the largest components satisfies $X_1(t) \to 1$ as $t \to \infty$ and

$$X_1(t) \to 0 \text{ as } t \to -\infty \qquad (*)$$

One might think that such a process was essentially unique, but in fact there is a highly non-trivial <u>entrance boundary</u> at time $-\infty$, i.e. many versions of the additive coalescent satisfying (*).