The birth of the strong components

Élie de Panafieu, **Sergey Dovgal**, Dimbinaina Ralaivaosaona,
Vonjy Rasendrahasina, and Stephan Wagner

(an upcoming work)

Seminar Plateau de Saclay, 17/06/2020

First cycles '1988

The First Cycles in an Evolving Graph

PHILIPPE FLAJOLET, DONALD E. KNUTH AND BORIS PITTEL

The purpose of this paper is to introduce analytical methods by which such questions can be answered systematically. In particular, we will apply the ideas to an interesting question posed by Paul Erdős and communicated by Edgar Palmer to the 1985 Seminar on Random Graphs in Posnań: "What is the expected length of the first cycle in an evolving graph?" The answer turns out to be rather surprising: The first cycle has length $Kn^{1/6} + O(n^{1/8})$ on the average, where

$$K = \frac{1}{\sqrt{8\pi}\, i} \int_{-\infty}^{\infty} \int_{1-i\infty}^{1+i\infty} e^{(\mu+2s)(\mu-s)^2/6} \frac{ds}{s}\, d\mu \approx 2.0337\,.$$

The form of this result suggests that the expected behavior may be quite difficult to derive using techniques that do not use contour integration.

Directed graphs '2020+

[de Panafieu, D., Ralaivaosaona, Rasendrahasina, Wagner]

Consider a random digraph from $\mathbb{D}(n, p)$ with n vertices, where each edge is drawn independently with probability p and is assigned a random direction (**Gilbert's model**).

- What is the probability that a digraph $\mathbb{D}(n, \frac{1}{n})$ is acyclic?

$$(2n)^{-1/3} e^{3/2} \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{1}{\mathrm{Ai}(-2^{1/3} s)} \mathrm{d}s \approx n^{-1/3} \cdot 2.19037\ldots$$

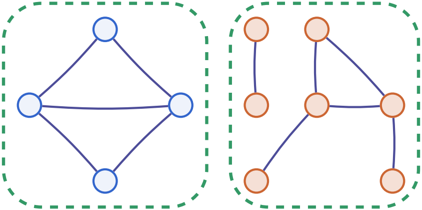
- What is the probability that the strongly connected components of a random digraph $\mathbb{D}(n, \frac{1}{n})$ are isolated vertices or cycles?

$$-2^{-2/3} \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{1}{\mathrm{Ai}'(-2^{1/3} s)} \mathrm{d}s \approx 0.69968786651_{+\mathcal{O}(n^{-1/3})}$$

Part I. Back to the origin: generating functions

The cartesian product

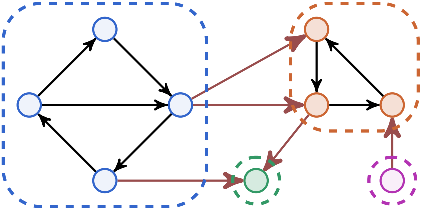
$$\left(a_0 + a_1\frac{z}{1!} + a_2\frac{z^2}{2!} + \dots\right)\left(b_0 + b_1\frac{z}{1!} + b_2\frac{z^2}{2!} + \dots\right) = c_0 + c_1\frac{z}{1!} + c_2\frac{z^2}{2!} + \dots$$

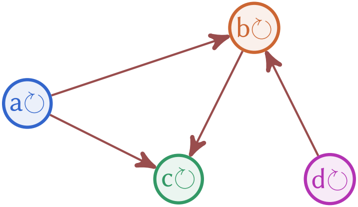


The convolution rule corresponding to EGF:

$$c_n = \sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k}$$

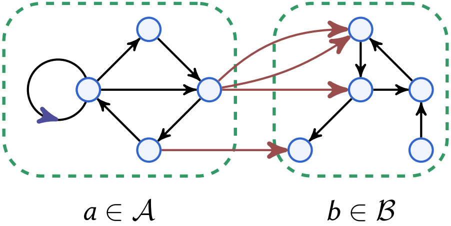
Directed graphs and their components





- Components a↻ , b↻ , c↻ and d↻ are *strongly-connected* components.
- Components a↻ and d↻ are *source-like* components
- Component c↻ is a *sink-like* component

The arrow product



The graphic generating function (GGF)

Let $\mathcal{F}$ be a family of digraphs and $D \in \mathcal{F}$. Let $n(D)$ denote the number of vertices, and $m(D)$ the number of edges of D.

Their EGF $F(z, w)$ and GGF $\widehat{F}(z, w)$ are defined as

$$F(z, w) := \sum_{D \in \mathcal{F}} \frac{z^{n(D)}}{n(D)!} \frac{w^{m(D)}}{m(D)!}, \quad \widehat{F}(z, w) := \sum_{D \in \mathcal{F}} e^{-n(D)^2 w/2} \frac{z^{n(D)}}{n(D)!} \frac{w^{m(D)}}{m(D)!}$$

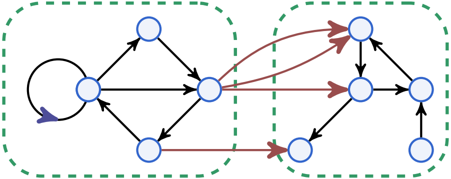
Proposition.

- $\widehat{F}(z, w) = \frac{1}{\sqrt{2\pi w}} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2w}\right) F(ze^{-ix}, w) \mathrm{d}x.$
- Gilbert's random model $\mathbb{P}_{n,p}$ is equidistributed to a Boltzmann distribution with parameter $\lambda = pn$ on the set of all multidigraphs.

$$\mathbb{P}_{n,p}(D \in \mathcal{F}) = e^{-pn^2/2} n! [z^n] \widehat{F}(z, p)$$

The graphic convolution product

$$\left(\sum_{n\geqslant 0} a_n(w)e^{-n^2w/2}\frac{z^n}{n!}\right)\left(\sum_{n\geqslant 0} b_n(w)e^{-n^2w/2}\frac{z^n}{n!}\right) = \sum_{n\geqslant 0} c_n(w)e^{-n^2w/2}\frac{z^n}{n!}$$



$a \in \mathcal{A}$ $b \in \mathcal{B}$

The convolution rule corresponding to GGF:

$$c_n(w) = \sum_{k+\ell=n}\binom{n}{k}(e^w)^{k\ell}a_k(w)b_\ell(w).$$

Part II. Families of directed graphs and their generating functions

The main enumeration theorem

Let $\mathcal{S}$ be a family of strongly connected digraphs, and let $\mathcal{D}_{\mathcal{S}}$ be the family of digraphs whose components are constrained to $\mathcal{S}$.

Theorem. GGF of $\mathcal{D}_{\mathcal{S}}$ is given by

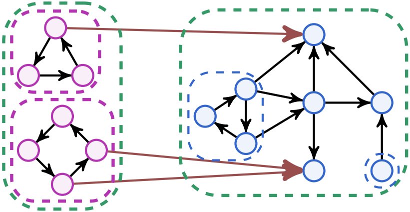
$$\widehat{D}_{\mathcal{S}}(z, w) = \frac{1}{\frac{1}{\sqrt{2\pi w}} \int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2w} - S(ze^{-ix}, w)\right) \mathrm{d}x}.$$

Moreover, if u marks the source-like components, then

$$\widehat{D}_{\mathcal{S}}(z, w, u) = \frac{\frac{1}{\sqrt{2\pi w}} \int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2w} + (u-1)S(ze^{-ix}, w)\right) \mathrm{d}x}{\frac{1}{\sqrt{2\pi w}} \int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2w} - S(ze^{-ix}, w)\right) \mathrm{d}x}.$$

Proof of the main enumeration theorem

Proof. $\mathcal{D}_\mathcal{S}$ with distinguished source-like components is an arrow product of a set of strong components and $\mathcal{D}_\mathcal{S}$.



are from distinguished source-like components.

are from the usual source-like components.

- $\widehat{D}_\mathcal{S}(z, w, u + 1) = \widehat{e^{uS(z,w)}} \cdot \widehat{D}_\mathcal{S}(z, w, 1)$.
- By letting $u = -1$, we obtain $\widehat{D}_\mathcal{S}(z, w) = \dfrac{1}{\widehat{e^{-S(z,w)}}}$.
- By plugging $u \mapsto u - 1$, we obtain $\widehat{D}_\mathcal{S}(z, w, u) = \dfrac{\widehat{e^{(u-1)S(z,w)}}}{\widehat{e^{-S(z,w)}}}$.

DAGs and elementary digraphs

Application 1. In DAGs, the only possible strong components are isolated vertices, $S(z, w) = z$.

$$\widehat{D}_{DAG}(z, w) = \frac{1}{\frac{1}{\sqrt{2\pi w}} \int\limits_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2w} - ze^{-ix}\right) \mathrm{d}x}.$$

Application 2. The *elementary digraphs* are those whose strong components are isolated vertices or cycles, $S(z, w) = z + \ln \frac{1}{1-zw}$.

$$\widehat{D}_{elem}(z, w) = \frac{1}{\frac{1}{\sqrt{2\pi w}} \int\limits_{-\infty}^{\infty} \frac{1 - zwe^{-ix}}{\exp\left(\frac{x^2}{2w} + ze^{-ix}\right)} \mathrm{d}x}.$$

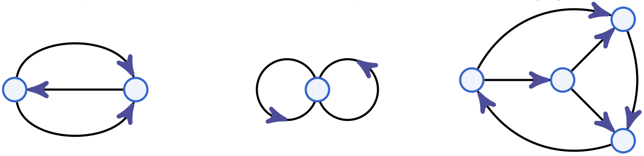
Complex components

The Birth of the Giant Component

Dedicated to Paul Erdős on his 80th birthday

Svante Janson, Donald E. Knuth, Tomasz Łuczak, and Boris Pittel

Is there a simple recurrence governing the leading coefficients $s_{10}, s_{20}, s_{30}, \ldots$, perhaps analogous to the relation we observed for ordinary connected components in (8.5)?



The EGF of strong components of *excess r* is

$$\mathrm{Strong}_r(z, w) = s_r w^r \frac{(zw)^{2r}}{(1 - zw)^{3r}} + w^r \frac{Q_r(zw)}{(1 - zw)^{3r-1}}.$$

$$(s_r)_{r=1}^{\infty} = \left(\frac{1}{2}, \frac{17}{8}, \frac{275}{12}, \frac{26141}{64}, \frac{1630711}{160}, \ldots \right).$$

Elementary digraphs with one bicyclic component

Application 3. Let $\widehat{H}_{\mathrm{bicycle}}$ be the GGF of elementary digraphs with one bicyclic component. Then,

$$\widehat{H}_{\mathrm{bicycle}}(z,w) \sim \frac{\dfrac{1}{\sqrt{2\pi w}} \displaystyle\int_{-\infty}^{\infty} \frac{1}{2} \frac{w^3 z^2 e^{-2ix}}{(1 - zwe^{-ix})^2} e^{-\frac{x^2}{2w} - ze^{-ix}} \mathrm{d}x}{\left(\dfrac{1}{\sqrt{2\pi w}} \displaystyle\int_{-\infty}^{\infty} (1 - zwe^{-ix}) e^{-\frac{x^2}{2w} - ze^{-ix}} \mathrm{d}x \right)^2}.$$

Proof. Apply the enumeration theorem with

$$S(z, w, \mathbf{v}) := z + \ln \frac{1}{1 - zw} + \mathbf{v} \cdot S_{\mathrm{bicycle}}(z, w),$$

where

$$S_{\mathrm{bicycle}}(z, w) = \frac{1}{2} \left(\frac{w^3 z^2}{(1 - zw)^3} + \frac{w^2 z}{(1 - zw)^2} \right)$$

and extract $[\mathbf{v}^1]$.

Source-like complex component

Generalised enumeration theorem. Let $\mathcal{S}$ and $\mathcal{H}$ be two disjoint families of strongly connected digraphs, and let $\mathcal{D}_{\mathcal{S},\mathcal{H}}$ be the family of digraphs whose components are contrained to $\mathcal{S}$ and $\mathcal{H}$. Let $\mathbf{u}$ and $\mathbf{v}$ mark source-like components from $\mathcal{S}$ and $\mathcal{H}$. Then,

$$\widehat{D_{\mathcal{S},\mathcal{H}}}(z, w, \mathbf{u}, \mathbf{v}) = \frac{\displaystyle\int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2w} + (\mathbf{u} - 1)S(ze^{-ix}, w) + (\mathbf{v} - 1)H(ze^{-ix}, w)\right) \mathrm{d}x}{\displaystyle\int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2w} - S(ze^{-ix}, w) - H(ze^{-ix}, w)\right) \mathrm{d}x}$$

Application 4. GGF of elementary digraphs with one source-like complex component from $\mathcal{S}$ is

$$W_{\mathcal{S}}(z, w) = \frac{\displaystyle\int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2w}\right) S(ze^{-ix}, w)\mathrm{d}x}{\displaystyle\int_{-\infty}^{+\infty} (1 - zwe^{-ix}) \exp\left(-\frac{x^2}{2w} - ze^{-ix} - S(ze^{-ix}, w)\right) \mathrm{d}x}$$

Proof. Take $H(z, w) = z + \ln \frac{1}{1-zw}$. Put $\mathbf{v} = 1$ and extract $[\mathbf{u}^1]$.

Part III. Asymptotic analysis

Asymptotic analysis: general scheme

- The probabilities of interest can be expressed as

$$\mathbb{P}_{n,p}(D \in \mathcal{F}) = e^{-pn^2/2} n! [z^n] \widehat{F}(z, p).$$

- $[z^n]\widehat{F}(z,p) = \dfrac{1}{2\pi i} \displaystyle\oint_{|z|=R} \dfrac{\widehat{F}(z,p)}{z^{n+1}} \mathrm{d}z.$
- For a given value of $p \to 0^+$, and for z fixed, find an asymptotic approximation of $\widehat{F}(z, p)$.
  - $\widehat{F}(z, p)$ is a product of integrals itself, each integral over $\mathbb{R}$.
  - Change the contour: preserve the starting and the finishing points, but let it pass through $x = x_0 \in i\mathbb{R}$ in the middle.
  - The dominant contribution is around $x = x_0 + \varepsilon$.

Horizontal Path. Path Γ.

- Dominant part of $[z^n]\widehat{F}(z, p)$ is when z is around $R \pm 0i$.

Asymptotics of the deformed exponent

Let $T(zw)$ and $U(zw)$ be the EGF of rooted and unrooted trees.

$$\phi(z,w) := \frac{1}{\sqrt{2\pi w}} \int_{-\infty}^{+\infty} \exp\left(-\frac{x^2}{2w} - ze^{-ix}\right) \mathrm{d}x = \sum_{n \geqslant 0} e^{-n^2 w/2} \frac{(-z)^n}{n!}.$$

(a) If $zw \in [0, e^{-1})$, then $\phi(z,w) \sim \dfrac{e^{-U(zw)/w}}{\sqrt{1 - T(zw)}}$

(b) If $1 - ezw = \theta w^{2/3}$, $\theta \to \infty$, then

$$\phi(z,w) \sim (2\theta)^{-1/4} w^{-1/6} \exp\left(-\frac{1}{2w} + \frac{\theta}{w^{1/3}} - \frac{2^{3/2}}{3}\theta^{3/2}\right)$$

(c) If $1 - ezw = \theta w^{2/3}$, $\theta \in \mathbb{C}$, then

$$\phi(z,w) \sim 2^{5/6} \pi^{1/2} w^{-1/6} \mathrm{Ai}(2^{1/3}\theta) \exp\left(-\frac{1}{2w} + \frac{\theta}{w^{1/3}}\right).$$

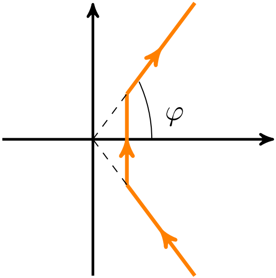
Generalised Airy function

The Airy function satisfies a linear differential equation

$$\mathrm{Ai}(z)'' - z\mathrm{Ai}(z) = 0$$

It can be expressed as an integral and its derivatives as well

$$\partial_z^r \mathrm{Ai}(z) = \frac{(-1)^r}{2\pi i} \int_{-i\infty}^{+i\infty} t^r \exp(-zt + t^3/3)\mathrm{d}t.$$



It is natural to extend this definition, so that $r \in \mathbb{Z}$ and deform the contour a little bit:

$$\mathrm{Ai}(r; z) := \frac{(-1)^r}{2\pi i} \int_{t \in \Pi(\varphi)} t^r \exp(-zt + t^3/3)\mathrm{d}t.$$

Generalised deformed exponent

Let $T(zw)$ and $U(zw)$ be the EGF of rooted and unrooted trees.

$$\psi_r(z,w) := \frac{1}{\sqrt{2\pi w}} \int_{-\infty}^{+\infty} (1 - zwe^{-ix})^r \exp\left(-\frac{x^2}{2w} - ze^{-ix}\right) \mathrm{d}x.$$

(a) If $zw \in [0, e^{-1})$, then $\psi_r(z,w) \sim e^{-U(zw)/w}(1 - T(zw))^{r-1/2}$

(b) If $1 - ezw = \theta w^{2/3}$, $\theta \to \infty$, then

$$\psi_r(z,w) \sim (2\theta)^{r/2-1/4} w^{r/3-1/6} \exp\left(-\frac{1}{2w} + \frac{\theta}{w^{1/3}} - \frac{2^{3/2}}{3}\theta^{3/2}\right)$$

(c) If $1 - ezw = \theta w^{2/3}$, $\theta \in \mathbb{C}$, then

$$\psi_r(z,w) \sim C \cdot D^r \cdot w^{-1/6+r/3} \mathrm{Ai}(r; 2^{1/3}\theta) \exp\left(-\frac{1}{2w} + \frac{\theta}{w^{1/3}}\right).$$

Computing the asymptotic probabilities

Theorem. In the multidigraph model, when $p = \frac{1}{n}(1 + \mu n^{-1/3})$,

- $\mathbb{P}_{n,p}(D_{n,p} \text{ is acyclic}) \sim (2n)^{-1/3} \cdot \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{e^{-\mu s - \mu^3/6}}{\mathrm{Ai}(-2^{1/3} s)} \mathrm{d}s$
- $\mathbb{P}_{n,p}(D_{n,p} \text{ is elementary}) \sim -2^{-2/3} \cdot \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{e^{-\mu s - \mu^3/6}}{\mathrm{Ai}'(-2^{1/3} s)} \mathrm{d}s$

The probability to have one complex component of excess r is asymptotically equal to

$$\mathbb{P}_{n,p}(\cdot) \sim s_r \cdot C \cdot D^r \cdot \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{\mathrm{Ai}(1 - 3r; -2^{1/3} s)}{(\mathrm{Ai}'(-2^{1/3} s))^2} e^{-\mu s - \mu^3/6} \mathrm{d}s.$$

Outside the critical window

- When $p = \lambda n^{-1}$, $\lambda < 1$, the probabilities can be obtained by applying large powers theorem to

$$\psi_r(z, w) \sim e^{-U(zw)/w}(1 - T(zw))^{r-1/2}$$

  for $zw < e^{-1}$.
- When $p = \lambda n^{-1}$, $\lambda > 1$, the knowledge of the roots of $\psi_r(z, w)$ is sufficient.
- When $p = n^{-1}(1 + \mu n^{-1/3})$, and $\mu \to -\infty$, we can apply semi-large powers theorem.

$$\mathbb{P}_{n,p}(D_{n,p} \text{ is elementary}) \sim 1 - \frac{1}{2|\mu|^3} + \mathcal{O}(|\mu|^{-6})$$

Part IV. Instead of the Post-Scriptum. The elusive coefficients s_r.

A few more theorems

Theorem. Let $p = n^{-1}(1 + \mu n^{-1/3})$. The probability that there are only bicyclic complex components (each weighted with u), is

$$\mathbb{P}_{n,p} \sim \frac{C}{2\pi i} \int_{-i\infty}^{i\infty} \frac{1}{\int_{\Pi(\varphi)} t e^{-2^{1/3} st + t^3/3} \exp\left(\frac{-1}{4t^3} u\right) \mathrm{d}t} \mathrm{d}s.$$

More generally, if multicyclic components are allowed, each marked with u_r corresponding to an excess r, the series will be

$$\mathbb{P}_{n,p} \sim \frac{C}{2\pi i} \int_{-i\infty}^{i\infty} \frac{1}{\int_{\Pi(\varphi)} t e^{-2^{1/3} st + t^3/3 - V(u_1, u_2, \ldots; t)} \mathrm{d}t} \mathrm{d}s,$$

where

$$V(u_1, u_2, \ldots; t) := \frac{s_1 u_1}{2t^3} + \frac{s_2 u_2}{4t^6} + \frac{s_3 u_3}{8t^9} + \ldots$$

Bootstrapping s_r

Let $p = n^{-1}(1 + \mu n^{-1/3})$. The coefficients s_r can be bootstrapped by considering asymptotic expansions in powers of μ^{-3} around $\mu \to -\infty$.

- First, forbid the complex components at all. We obtain

$$\mathbb{P}_{n,p}(D_{n,p} \text{ is elementary}) \sim 1 - \frac{1}{2|\mu|^3} + \frac{?}{|\mu|^6} + \dots$$

- Then, the probability of having only bicyclic components is

$$\mathbb{P}_{n,p}(D_{n,p} \text{ has only bicyclic c.c.}) \sim s_r \left(\frac{1}{|\mu|^3} + \frac{?}{|\mu|^6} + \dots \right)$$

- Adding the component of excess r, and summing up the coefficients at $|\mu|^{-3r}$, we obtain the sequence.

Conclusion

Conclusion

1. The phase transition curves for DAG, elementary digraphs and analysis of complex components can be finally completed.
2. The technique is highly flexible with respect to different digraph models (with or without loops or 2-cycles)
3. Still a lot of questions open (and probably doable!):
    - Statistics of random DAGs (sinks, sources)
    - Asymptotics of strongly connected graphs
    - Simultaneous asymptotics of sink-like and source-like components
    - Cubic kernels (digraphs)
    - Digraphs with degree contraints
    - Giant component of a digraph
    - Triple, quadruple arrow product?
    - Analysis of 2-SAT with similar level of precision
    - …(enough for a PhD thesis or so) …

Thank you for your attention.

Bonus 1. Saddle point analysis

The target generalised integral is given by

$$I = \int h(x_0 + t)^r e^{f(x_0+t)} \mathrm{d}t,$$

where

$$f(x) = -\frac{x^2}{2w} - ze^{-ix}, \quad h(x) = 1 - zwe^{-ix}.$$

The stationary point is defined by relation

$$f'(x) = 0 \quad \Leftrightarrow \quad x_0 = iT(zw).$$

The second derivative of $f(x)$ vanishes when $zw = e^{-1}$ which also corresponds to $p = n^{-1}$. The limiting stationary point is $x_0 = i$.

Bonus 2. Simple digraphs and multidigraphs

The basic deformed exponent corresponding to simple digraphs is

$$\phi^{(simple)}(z, w) = \phi(z\sqrt{1 + w}, \log(1 + w)).$$

Elementary digraphs can be adjusted by forbidding loops and 2-cycles which yields

$$S(z, w) = z + \ln \frac{1}{1 - zw} - zw - \frac{(zw)^2}{2}.$$

All the obtained asymptotic approximations can be readily used to obtain the asymptotics of simple digraphs directly.

Bonus 3. Product decomposition

The denominator of the GGF of DAGs can be expressed in terms of the EGF of all the graphs.

$$\phi(z,w) = \sum_{n \geqslant 0} e^{-n^2 w/2} \frac{(-z)^n}{n!} = MG(-z,-w)$$
$$= e^{-U(zw)/w + V(zw)} \sum_{k \geqslant 0} \mathrm{Complex}_k(zw)(-w)^k$$

This allows to express both the asymptotic of DAGs and elementary digraphs as an infinite sum, because

$$D_{DAG}(z,w) = \frac{1}{MG(-z,-w)} \quad \text{and} \quad D_{\mathrm{elem}}(z,w) = \frac{1}{MG(-z,-w) + zw\partial_z MG(-z,-w)}.$$