Anarchy in the Aggregation Equation

Cameron Davies

June 2, 2021

1 Introduction to the Aggregation Equation and Attractive-Repulsive Potentials

For the purposes of this talk, we will consider the aggregation equation, which is the following PDE in μ:

$$\frac{\partial \mu}{\partial t} = \nabla \cdot (\mu \nabla W * \mu).$$

Solutions to this equation take the form of parametrized families $\mu(t)$ of Borel probability measures ($\mathcal{P}(\mathbb{R}^n)$). We think of $W : \mathbb{R}^n \to \mathbb{R}$ as a fixed interaction potential, and the choice of W determines the behaviour of solutions to the aggregation equation. The aggregation equation has been used to model many natural phenomena in areas ranging from animal swarming, to chemotaxis, to nanomaterials, to game theory (references in [1]).

We can associate an interaction energy to the aggregation equation:

$$\mathcal{E}_W(\mu) := \frac{1}{2} \iint_{\mathbb{R}^n \times \mathbb{R}^n} W(x-y) d\mu(x) d\mu(y).$$

This serves as a Lyapunov functional for the aggregation equation; local minimizers of the energy correspond to steady states of the aggregation equation. Additionally, if you're familiar with Wasserstein distances and the associated gradient flows (see [3] for a reference), the interaction energy can be realized as the Wassersterin d_2 gradient flow of this interaction energy. Thus, given these fundamental connections to the aggregation equation, we'll study the aggregation equation by studying both local and global energy minimizers of this energy. Moreover, we notice that this energy is translation invariant in μ, so we will often consider it as a functional on the space $\mathcal{P}_0(\mathbb{R}^n)$ of centred probability measures, defined by:

$$\mathcal{P}_0(\mathbb{R}^n) := \{\mu \in \mathcal{P}(\mathbb{R}^n) | \int x d\mu(x) = 0\}$$

In particular, we will consider radially symmetric 'power law' interaction potentials

$$W_{\alpha,\beta}(x) = \frac{|x|^\alpha}{\alpha} - \frac{|x|^\beta}{\beta}$$

for real numbers $\alpha > \beta \geq 2$. Such interaction potentials describe how pairs of particles interact, and have been normalized to have minimum at $|x| = 1$. If $\mu = m\delta_{x_1} + (1-m)\delta_{x_2}$, then

$$\mathcal{E}_{W_{\alpha,\beta}}(\mu) = m(1-m) \left[\frac{|x_1 - x_2|^\alpha}{\alpha} - \frac{|x_1 - x_2|^\beta}{\beta} \right].$$

The qualitative behaviour of $W_{\alpha,\beta}$ depends heavily on the values of $\alpha > \beta$, as the following figure illustrates, where we have graphed $W_{\alpha,\beta}$ against $|x|$:

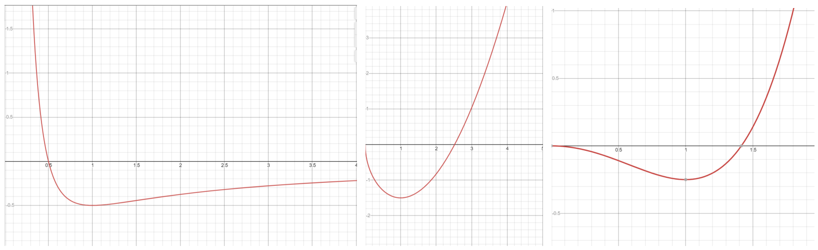


Figure 1: $(\alpha, \beta) = (-1, -2)$, $(2, 0.5)$, and $(4, 2)$, respectively.

The leftmost figure, where $(\alpha, \beta) = (-1, -2)$ is in what's called the **strongly repulsive** regime, where $0 > \alpha > \beta$. In this regime, the potential dictates that it is infinitely expensive to have *any* mass concentrated at a single point. The middle figure represents an intermediate regime, where the potential ensures that it is possible to have mass concentrated at a point, but where the energy (at least in most cases) can be decreased rapidly by spreading out point masses a bit. Finally, the rightmost picture is a typical example of a potential in the **mildly repulsive** regime, $\alpha > \beta \geq 2$. In this regime, the potential does still have a local maximum at 0, but its rate of change near 0 is small enough to ensure that mass can (quite often) concentrate.

The behaviour of the system depends on the values of α and β. For high values of α, it becomes 'expensive' for any mass to be separated by distance more than 1. However, it becomes relatively 'cheaper' to have mass at distance exactly 1 apart. Both of these phenomena are illustrated in the following Desmos Graph. Using this intuition, Lim and McCann [1] showed that, for sufficiently large α, the global minimizer of the interaction energy is given by the uniform distribution over the vertices of a unit simplex. In $\mathbb{R}^n$, the vertices of a unit simplex can be thought of as a collection of $n + 1$ points, each lying at Euclidean distance 1 from each other.

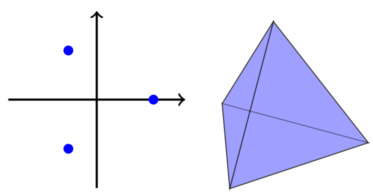


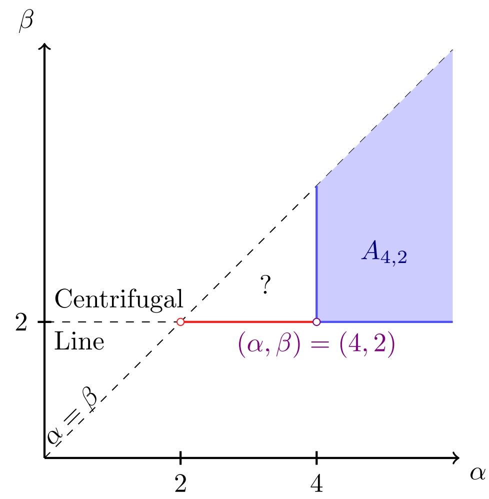
Figure 2: Simplices in two and three dimensions. Notice that, in three dimensions, the minimizers we concern ourselves with distribute their mass uniformly over the vertices of the simplex, and not the entire simplex.

2 Anarchy in the Aggregation Equation

In recent (not yet published) work, Tongseok Lim, Robert McCann, and I characterized minimizers of the interaction energy for a large subset of the mildly repulsive regime. Our results differ somewhat for the cases $n = 1$ and $n \geq 2$, and hence, in what remains of this talk, we will focus on the $n \geq 2$ case. In particular, if $\beta = 2$ and $\alpha \in (2, 4)$ (red), the minimzer is uniquely given by a spherical shell. If (α, β) lie in the blue regime

$$A_{4,2} = \{(\alpha, \beta) \in \mathbb{R}^2 \mid \alpha > \beta, \alpha \geq 4, \beta \geq 2, (\alpha, \beta) \neq (4, 2)\},$$

then the unique minimizer is given by the uniform distribution over the vertices of a unit $n-$simplex. These regimes are summarized in the diagram below:



When $(\alpha, \beta) = (4, 2)$, minimizers are non-unique. Perhaps this is to be expected, given that this point lies at the border of the regimes of where the interaction energy is minimized by a spherical shell and the uniform distribution over the vertices of a unit $n-$simplex, respectively. However, there wind up being many more minimizers than these two at $(\alpha, \beta) = (4, 2)$. but are characterized by their second moment tensor, $\int x \otimes x d\mu(x)$:

Theorem 1 (Theorem 1 (Davies-Lim-McCann)) *$\mu \in \mathcal{P}_0(\mathbb{R}^n)$ is a minimizer of $\mathcal{E}_{W_{4,2}}$ if and only if μ is concentrated on the centered sphere of radius $\sqrt{\frac{n}{2n+2}}$ with*

$$\int x \otimes x \, d\mu(x) = \left(\int x_i x_j d\mu(x) \right)_{1 \le i,j \le n} = \frac{1}{2n+2} \mathrm{Id}.$$

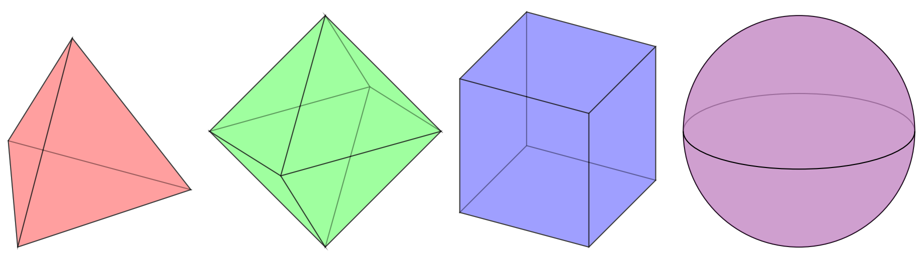


Figure 3: Examples of Minimizers of $\mathcal{E}_{W_{4,2}}$, each of which is uniformly distributed over the vertices/extreme points of each of these shapes.

3 Proof of Theorem 1

Proof Strategy: The proof proceeds in four steps, which are included (to various degrees) in these notes:

- **Step 1:** Show that all minimizers have the same moment of inertia tensor, $\lambda \mathrm{Id}$.
- **Step 2:** Show that all minimizers are concentrated on a spherical shell of radius $r = \sqrt{n\lambda}$
- **Step 3:** Show that any measure μ concentrated on a shell of radius r with $I(\mu) = \lambda \mathrm{Id}$ is a minimizer.
- **Step 4:** Calculate λ (simple calculation; omitted)

Proof of Step 1: First, we show that all minimizers must have the same moment of inertia tensor using a convexity argument. Since convex combinations of measures in $\mathcal{P}_0(\mathbb{R}^n)$ are also in $\mathcal{P}_0(\mathbb{R}^n)$, we take candidate minimizers μ_0 and μ_1 and consider the interaction energy evaluated at a convex linear combination of them:

$$\mathcal{E}_{W_{4,2}}(\mu_0 + t(\mu_1 - \mu_0)) = \mathcal{E}_{W_{4,2}}(\mu_0) + 2t \iint_{\mathbb{R}^n \times \mathbb{R}^n} W_{4,2}(x-y) d\mu_0(x) d(\mu_1 - \mu_0)(y) + t^2 \mathcal{E}_{W_{4,2}}(\mu_1 - \mu_0).$$

Clearly, this is a quadratic function in t, so we compute the second derivative, $2\mathcal{E}_{W_{4,2}}(\mu_1-\mu_0)$. In particular, if the second derivative is positive, then $\mu_0+t(\mu_1-\mu_0)$ has strictly lower interaction energy than either μ_0 or μ_1, a contradiction to our hypothesis that μ_0 and μ_1 are minimizers.

Since $\int d(\mu_1-\mu_0)=\int x d(\mu_1-\mu_0)=0$, write $|x-y|^2=(x-y)\cdot(x-y)$ and $|x-y|^2=((x-y)\cdot(x-y))^2$ and expand the dot product to see that:

$$2\mathcal{E}_{W_{4,2}}(\mu_1-\mu_0)=\iint\limits_{\mathbb{R}^n\times\mathbb{R}^n}\frac{|x-y|^4}{4}-\frac{|x-y|^2}{2}d(\mu_1-\mu_0)(x)d(\mu_1-\mu_0)(y)$$

$$=\frac{1}{2}\iint\limits_{\mathbb{R}^n\times\mathbb{R}^n}|x|^2|y|^2d(\mu_1-\mu_0)(x)d(\mu_1-\mu_0)(y)+\iint\limits_{\mathbb{R}^n\times\mathbb{R}^n}(x\cdot y)^2d(\mu_1-\mu_0)(x)d(\mu_1-\mu_0)(y)$$

$$=\frac{1}{2}(\mathrm{Tr}I(\mu_1-\mu_0))^2+\mathrm{Tr}I(\mu_1-\mu_0)^2.$$

This is a sum of squares and, in particular, the second term is the sum of the squares of each of the entries of the second moment tensor $I(\mu_1-\mu_0)$. Hence, $\frac{d^2}{dt^2}\mathcal{E}_{W_{\alpha,\beta}}(\mu_0+t(\mu_1-\mu_0))$ is equal to zero if and only if $I(\mu_1)=I(\mu_0)$. In particular, this means that μ_0 and μ_1 can only be minimizers if they share the same second moment tensor - i.e. all minimizers necessarily have the same second moment tensor.

To show that the common second moment tensor takes the form $\lambda\mathrm{Id}$, we extended an omitted convexity argument of Orlando Lopes [2] to show that that $\mathcal{E}_{W_{4,2}}$ is convex. Hence, there exists at least one spherically symmetric minimizer of $\mathcal{E}_{W_{4,2}}$ (averaging a minimizer over its rotations also yields a minimizer). Using this spherically symmetric minimizer, we see that the off-diagonal terms of the second moment tensor

$$\int x\otimes x\,d\mu(x)=\left(\int x_ix_jd\mu(x)\right)_{1\leq i,j\leq n}$$

are zero and the diagonal terms are all the same constant λ.

Proof of Step 2: By an analogous calculation to Step 1, using that $\int d\mu=1$ whereas $\int d(\mu_1-\mu_0)=0$, we see that if μ is a centred probability measure then,

$$2\mathcal{E}_{W_{4,2}}(\mu)=\frac{1}{2}\int|x|^4d\mu(x)+\frac{1}{2}(\mathrm{Tr}I(\mu))^2+\mathrm{Tr}I(\mu)^2-\mathrm{Tr}I(\mu).$$

In particular, if $I(\mu)=\lambda\mathrm{Id}$, then

$$2\mathcal{E}_{W_{4,2}}(\mu)=\frac{1}{2}\int|x|^4d\mu(x)+\frac{1}{2}n^2\lambda^2+n\lambda^2-n\lambda.$$

Thus, working with the optimal λ, which is to be determined, we see that μ is an actual minimizer of $\mathcal{E}_{W_{4,2}}(\mu)$ if and only if it minimizes $\int|x|^4d\mu(x)$. Although I do not have time to go through the details, a bit of introductory optimal transport, applied in conjunction with Jensen's inequality shows that $\int|x|^4d\mu(x)$ is minimized over all measures μ with $\mathrm{Tr}I(\mu)=n\lambda$ precisely if μ is concentrated on the sphere σ of radius $r=\sqrt{n\lambda}$. This concludes the proof of Step 2.

Proof of Step 3: Thus, since we know there's a spherically symmetric minimizer, the requirement that the minimizer concentrate on a spherical shell ensures that the spherical shell of radius r must minimize $\mathcal{E}_{W_{4,2}}$. And, since $\int x_1^2+...+x_n^2d\sigma(x)=n\lambda$, we can use symmetry to see that $I(\sigma)=\lambda\mathrm{Id}$.

Thus, we revisit our earlier formula for $\mathcal{E}_{W_{4,2}}$, which states that:

$$2\mathcal{E}_{W_{4,2}}(\mu)=\frac{1}{2}\int|x|^4d\mu(x)+\frac{1}{2}n^2\lambda^2+n\lambda^2-n\lambda.$$

Our spherical shell condition determines the first term in this expression, and our second moment condition totally determines the last three, ensuring sufficiency of these conditions.

Proof of Step 4: (Omitted) single variable optimization

References

[1] Tongseok Lim and Robert J. McCann. Isodiametry, variance, and regular simplices from particle interactions. To appear in *Archive Rational Mech. Analysis* (2021) https://doi.org/10.1007/s00205-021-01632-9

[2] Orlando Lopes. Uniqueness and radial symmetry of minimizers for a nonlocal variational problem. *Comm. Pure. Appl. Anal.* 18 (2019) 2265-2282.

[3] Filippo Santambrogio. *Optimal transport for applied mathematicians.* Birkhäuser/Springer, Cham, 2015.