COCOMAS: Miscellaneous notes

March 23, 2022

Contents

This document contains a brief summary of the research lines followed by the COCOMAS research project, including several partial, unpublished results. This research has been funded by the spanish MINECO (Ministerio de Ciencia e Innovación), grant ref. MTM2017-91054-EXP.

1 Introduction and preliminary considerations

An important problem in Aquatic Sciences is to determine the distribution of aquatic particles at a given environment. Particle sizes range over several orders of magnitude in this context, from the submicron scale (nanogels and colloids) to the macroscopic scale ("marine snow" aggregates) [6, 8, 45]. These particles are quite diverse and may include for instance phytoplankton cells, gels, fecal pellets, amorphous organic detritus and heavy elements. The particle size distribution is a fundamental tool in the study of aquatic ecosystems; its knowledge allows to understand such diverse things as the light penetration in the water

column, the demise of algal blooms, the iron cycling in the ocean or the carbon fluxes, see for instance the review papers [8, 26].

When it comes to the improvement of our understanding of the biogeochemical cycling and the fate of organisms in aquatic environments, one of the main concerns is to understand what is the global carbon budget. This is connected with Aquatic Sciences because the main sources of particulate organic carbon are ultimately biological; the ocean holds one of the largest stocks of organic carbon on earth indeed [5]. One among a number of interesting implications is the fact that during sedimentation processes heavy particles drag to the bottom of lakes and oceans a large amount of CO_2, thus getting trapped and leaving the carbon cycle. This contributes to the mitigation of global warming and has been already the subject of some scientific studies (see for instance [17, 41]).

It is therefore important to be able to formulate reliable models allowing to predict the time evolution of particle size distributions in several contexts within the realm of Aquatic Sciences. There are many physical, chemical and biological processes going on in aquatic environments that have an influence on the size distribution of aquatic particles at a given time. This complexity renders the design of useful models a formidable research challenge. In this project we take up this challenge from different points of view, using methodologies based on inverse problem solving.

2 Coagulation models in oceanography: mathematical properties

Aggregate formation is nowadays recognized to be one of the most influential processes on the overall dynamics of aquatic particles (see for example [8, 19, 34]). In fact, many particles in surface waters are in the form of small cells, which have slow sinking rates. Particle sinking rates depend on particle size: aggregation mechanisms are important here because they allow the repackaging of small particles into larger ones. This is known to be one of the main processes controlling the flux of particles from the surface to the deep ocean. Then aggregation mechanisms become crucial to determine the distribution of chemical elements (carbon, other nutrients like P, N, Si and trace elements as Th, Pb, Ag, Se) in lakes and oceans, both at the surface and at the bottom. They can also affect ecological interactions by, e.g., selectively removing some algal species or altering zooplankton feeding by both removing food and combining small particles into packages of eatable size [26]. Therefore, aggregation mechanisms are highly relevant for aquatic ecosystems and biogeochemical cycling in a broad sense.

A large number of quantitative models in this field are based on the pure coagulation equation (continuous version), which reads as follows:

$$\frac{\partial n(t,x)}{\partial t} = \frac{1}{2}\int_0^x K(x-y,y)n(t,x-y)n(t,y)\,dy - \int_0^\infty K(x,y)n(t,x)n(t,y)\,dy\,. \qquad (2.1)$$

This model describes aggregation phenomena by pairwise interactions between particles. More sophisticated interactions are assumed to be decomposable as a sequence of binary aggregation events. Here x accounts for a measure of "size", such that when a particle with size x meets a particle with size y, the resulting aggregate has size $x + y$. This is an

assumption on how the model works; e.g. when dealing with spherical particles, the correct notion of size is the usual volume (note that this can be modified if we want to account for particles having fractal dimension). The coagulation kernel $K(x,y)$ codifies how likely is the interaction $(x)+(y) \to (x+y)$. There are lots of possibilities, depending on the detailed knowledge we may have on the physical coagulation interactions.

Actually, in oceanographical contexts we want to consider the pure coagulation equation with a source-sink term S added:

$$\frac{\partial n(t,x)}{\partial t} = \frac{1}{2}\int_0^x K(x-y,y)n(t,x-y)n(t,y)\,dy - \int_0^\infty K(x,y)n(t,x)n(t,y)\,dy + S(x,n(t,x)). \tag{2.2}$$

Particle synthesis and settling are described by means of S; the specific choice

$$S(x,n(t,x)) = I(x) - \frac{w(x)}{Z}n(t,x) \tag{2.3}$$

is representative. Here the source term I depends only on particle size and has a net contribution only for small sizes. The sinking term is described in terms of the depth of the water layer, Z, and the settling speed, given by Stokes' law or some variants thereof [9, 27].

2.1 Quick reference on the pure coagulation equation

It is customary to assume that the coagulation kernel satisfies the following properties:

1. $K : \mathbb{R}_0^+ \times \mathbb{R}_0^+ \to \mathbb{R}_0^+$,
2. $K(x,y) = K(y,x)$ for every $x, y \geq 0$.

We will take for granted that the initial condition is such that

$$\int_0^\infty (1+x)n_0(x)\,dx < \infty.$$

Then we can ensure the existence of global solutions whenever the coagulation kernel satisfies, for some $C > 0$,

$$K(x,y) \leq C(1+x+y), \forall x, y \geq 0.$$

Otherwise gellation phenomena may take place (formation of infinitely-sized aggregates in finite time, see e.g. [20]); in this context we are not interested in those.

We introduce the moments

$$M_i(t) = \int_0^\infty x^i n(t,x)\,dx\,.$$

Given the above assumptions, M_1 (total mass) is preserved during evolution. The number of fragments decreases with time,

$$\dot{M}_0(t) = -\frac{1}{2}\int_0^\infty \int_0^\infty K(x,y)n(t,x)n(t,y)\,dxdy\,.$$

Under certain additional assumptions on the initial datum and on the kernel K the long-time behaviour is self-similar [24, 21]. In order to have such dynamics, we adopt the following homogeneity prescription:

$$\text{there exists some } \lambda \leq 1 \text{ such that} \quad K(\gamma x, \gamma y) = \gamma^{\lambda} K(x,y), \quad \forall \gamma > 0\,. \tag{2.4}$$

Thus,

$$K(x,y) = y^{\lambda} K(x/y, 1) \quad \text{and we define} \quad \kappa(z) := K(z,1)\,.$$

In this fashion,

$$K(x,y) = y^{\lambda}\kappa(x/y) = x^{\lambda}\kappa(y/x), \quad \text{in other words,} \quad \kappa(z) = z^{\lambda}\kappa(1/z)\,.$$

By self-similar behaviour, here we shall understand that there exists some profile g such that

$$n(t,x) = t^{2\nu} g(t^{\nu} x) \quad \text{or, more specifically,} \quad n(t,x) \simeq t^{2\nu} g(t^{\nu} x) \quad \text{for advanced times,}$$

for some $\nu \in \mathbb{R}$. In this way $M_1(n) = M_1(g)$ -other moments will not agree. Note here that there are more general scaling laws that may be needed to accommodate certain cases (borderline homogeneities, power laws in size, etc, see for instance [43]).

There are two specific cases where self-similar behavior can be checked explicitly:

- The constante kernel: we have the explicit solution $n(t,x) = 4\frac{e^{-2x/t}}{t^2}$. Actually, we have a one-parametric family of self-similar profiles $g(z) = \frac{4}{M_1} e^{-\frac{2z}{M_1}}$.
- The sum kernel: we have the explicit solution $n(t,x) = \frac{e^{-t}}{\sqrt{2\pi x^{3/2}}} \exp\{-e^{-2t}x/2\}$. Indeed, there is a two-parametric family of self-similar profiles $g(z) = \frac{\rho\sigma^{3/2}}{\sqrt{2\pi}} z^{-3/2} e^{-\frac{z}{2\sigma}}$; in this case the temporal scaling is not given by $s(t) = t^{\nu}$ but rather $s(t) = e^{2\rho t}$.

Well-posedness for (2.2) can be obtained as a variant of the results in e.g. [1, 2, 3, 18, 39]. A detailed proof for bounded kernels and sources can be found in [44]; the proof uses Picard iterants. To the best of our knowledge, there are no works concerning the dynamical behavior of (2.2)–(2.3).

2.2 Construction of steady states for (2.2)–(2.3)

Here we ellaborate on the problem of finding nontrivial solutions for

$$\frac{1}{2}\int_0^x K(x-y,y) n_\infty(x-y) n_\infty(y)\, dy - \int_0^\infty K(x,y) n_\infty(x) n_\infty(y)\, dy + S(x, n_\infty(x)) = 0, \tag{2.5}$$

at least for suitable S. Given that $I(x)$ does not vanish identically, note that $N = 0$ is never a steady state.

We currently lack an existence proof that is general enough to cover the existence of steady states in all the cases we would be interested in. Banach-fixed-point arguments on (2.5) seem to work only under smallness assumptions on the settling rate, which, besides,

needs to be separated from zero. For instance, we can show the following: given a constant settling speed $w(x) = w$ (resp. uniformly separated from zero), the condition $6\|K\|_\infty \|I\|_1 \leq w^2$ ensures that there is a stationary solution n_∞ of (2.5)–(2.3) such that $\|n_\infty\|_1 < \frac{w}{3\|K\|_\infty}$. In this case, if there exist more stationary solutions they must verify that $\|n_\infty\|_1 \geq \frac{w}{3\|K\|_\infty}$. However, none of the situations we would be interested in fall under this framework, and the mechanism used to prove this statement does not seem to extend much further.

Proofs using semigroup methods as those in [21, 28] would be more convenient (as convergence to the steady state also follows from the fixed-point argument), however, they do no extend easily precisely due to the fact that the settling speed is not separated from zero. Ad hoc techniques can be applied to some specific kernels, *e.g.* Laplace transform methods can be use to prove existence of steady states for the constant coagulation kernel.

It is also interesting to mention that the numerical simulations performed in [25] seem to indicate that the steady states, should they exist, may not resemble power-law profiles. As power laws give very convenient fitting functions to oceanographical size spectrum data -see *e.g* [7, 26], a more detailed investigation on the mathematical properties of the steady state solutions is needed.

3 Some comments on size-class models for aggregation in aquatic contexts

Part of the results sketched here were obtained in [23].

The models that were introduced in the previous section regard aggregate size as a continuous variable. In the parlance of oceanographers, those are *size spectrum models*. Sometimes such models are regarded as difficult to use, hence simpler descriptions are sought. Various families of *size class models* are also common in this field; these representations lump together a series of size ranges into a small number of representative size classes (e.g. it is typical to consider just two classes "small" and "large" particles). The discretization of the size range is usually performed on the basis of the apertures of various filters that can be used to separate particles larger than a given size from a given sample. This happens also in such a way that particle properties differ across classes; for instance, particles labelled as "small" are typically suspended in the water column, while those labelled as "large" gradually sink to the bottom.

3.1 A simple instance of size class models

A representative model under this paradigm is the one used in [19], where the temporal dynamics of PCHO (dissolved polissacarides) and TEP (transparent exopolymeric particles) are described during a phytoplankton bloom experiment (one of the first studies to point out the importance of assembly interactions in the formation of marine particles). A slight simplification of the aforementioned model -using a simpler source term-, reminiscent of the

variants in [6], reads

$$\begin{cases} \dfrac{dPCHO}{dt} = S - a\,PCHO^2 - b\,PCHO \cdot TEP \\ \dfrac{dTEP}{dt} = a\,PCHO^2 + b\,PCHO \cdot TEP - r_{sed} TEP \end{cases} \tag{3.1}$$

The parameters a, b, S and r_{sed} are all taken to be positive. For operational purposes, we think of PCHO as "small particles" and TEP as "large particles", see the discussion in the previous paragraph. The following interactions are described:

1 PCHO is synthetized by specific biochemical processes, at a rate given by S.

2 Two PCHO particles can aggregate to form a TEP particle (second order kinetics) at a rate given by a.

3 A PCHO particle may aggregate with a TEP particle to yield a TEP particle (second order kinetics) at a rate given by b.

4 TEP particles gradually sink to the bottom and thus leave the system by sedimentation effects, which take place at a rate r_{sed}.

This model has a unique equilibrium point in the first quadrant, given by

$$\mathrm{TEP}_\infty = \frac{S}{r_{sed}}, \quad \mathrm{PCHO}_\infty = \frac{-b\mathrm{TEP}_\infty + \sqrt{(b\mathrm{TEP}_\infty)^2 + Sa}}{2a}\,.$$

Using a standard phase plane analysis and Dulac's criterion, it can be shown that this equilibrium is a global attractor over the first quadrant [23].

The model (3.1) is appealing because of the neat balance

$$\frac{d}{dt}[PCHO + TEP] = S - r_{sed} TEP\,.$$

This is reminiscent of the fact that aggregation does not change the total mass of the system (only synthesis and sedimentation will). However, the equations are derived in a phenomenological way precisely in order to match this balance; note in particular that this model does not follow a natural chemical kinetics. To try to obtain more natural kinetics one could consider describing interactions *2* and *3* as

$$PCHO + PCHO \xrightarrow{a/2} 2TEP, \quad PCHO + TEP \xrightarrow{b} 2\,TEP\,,$$

or, alternatively,

$$PCHO + PCHO \xrightarrow{a} PCHO + TEP, \quad PCHO + TEP \xrightarrow{b} 2\,TEP\,.$$

This does not make full sense in either case, particularly regarding the kinetics for *2*. Perhaps a slightly more reasonable kinetics for interactions *2* and *3* would be

$$PCHO + PCHO \xrightarrow{a} TEP, \quad PCHO + TEP \xrightarrow{b} 2\,TEP\,. \tag{3.2}$$

This will render the alternative model

$$\begin{cases} \dfrac{dPCHO}{dt} = S - 2a\,PCHO^2 - b\,PCHO \cdot TEP \\ \dfrac{dTEP}{dt} = a\,PCHO^2 + b\,PCHO \cdot TEP - r_{sed} TEP\,. \end{cases} \tag{3.3}$$

Again, and with a very similar analysis, this model can be shown to have a unique equilibrium point in the first quadrant; this equilibrium is a global attractor over the first quadrant [23]. Therefore, this alternative proposal seems a bit more reasonable. In any case, note that we lose the nice balance relation for $PCHO + TEP$. It can also be argued that the chemical kinetics used to deduce it is not fully satisfactory anyway, should we have more precise information on the sizes of the interacting particles via *3*. This is a drawback that all size class models have, as coarse-graining the system into just a few categories disregards a huge amount of information.

Given that the former models cannot be constructed from first principles but are rather phenomenological, an identifiability analysis seems necessary before attempting any parameter-identification exercise. We ran identifiability tests using DAISY [4] to conclude that both models (3.1) and (3.3) are globally identifiable.

3.2 Variants and extensions of the model

The models in the previous section are designed to operate on small spatial volumes (order of cubic meters). Other studies may require to use models that span bodies of water that are way larger, e.g. certain regions of the Mediterranean sea. In this context, the former models would just consider a single depth layer and a system which is spatially homogeneous in latitude and longitude. Here we will briefly consider extensions that allow to include information about several depth layers. This looks reasonable once we work with oceanographic datasets collected at monitoring points that are located at the same latitude and longitude ("stations") but on different depths -say five or six different depths. Variants with respect to latitude and longitude may be accounted for by changing the parameters.

The model in section 3.1 can be easily extended to accommodate more than one depth layer. For instance, we can consider an upper layer (layer 1) and a lower layer beneath it (layer 2). Then, with obvious notations, we would consider four unknowns

$$PCHO_1(t), TEP_1(t), PCHO_2(t), TEP_2(t)$$

with their corresponding synthesis rates S_1, S_2, aggregation rates a_1, b_1, a_2, b_2 and sedimentation rates $t_{1,2}$ (transfer from the upper to the lower layer) and r_{sed} (transfer from the lower layer to trapped matter); all those are taken to be positive. The full model would

read

$$\begin{cases} \dfrac{dPCHO_1}{dt} = S_1 - a_1 PCHO_1^2 - b_1 PCHO_1 \cdot TEP_1 \\ \dfrac{dTEP_1}{dt} = a_1 PCHO_1^2 + b_1 PCHO_1 \cdot TEP_1 - t_{1,2} TEP_1 \\ \dfrac{dPCHO_2}{dt} = S_2 - a_2 PCHO_2^2 - b_2 PCHO_2 \cdot TEP_2 \\ \dfrac{dTEP_2}{dt} = a_2 PCHO_2^2 + b_2 PCHO_2 \cdot TEP_2 + t_{1,2} TEP_1 - r_{sed} TEP_2 \, . \end{cases}$$

Note that the upper layer is not affected by the dynamics of the lower layer at all. We easily find out that equilibrium points are characterized by

$$TEP_1 = \frac{S_1}{t_{1,2}}, \quad TEP_2 = \frac{S_2 + t_{12} TEP_1}{r_{sed}}.$$

This structure can be generalized to an arbitrary number of vertical layers and an ODE system results with the property that each layer is only affected by the ones it has above it and we may transfer the information layer by layer from top to bottom. For completeness we write down a model with $N \geq 3$ layers:

$$\begin{cases} \dfrac{dPCHO_1}{dt} = S_1 - a_1 PCHO_1^2 - b_1 PCHO_1 \cdot TEP_1 \\ \dfrac{dTEP_1}{dt} = a_1 PCHO_1^2 + b_1 PCHO_1 \cdot TEP_1 - t_{1,2} TEP_1 \\ \dfrac{dPCHO_i}{dt} = S_i - a_i PCHO_i^2 - b_i PCHO_i \cdot TEP_i, \; i = 2, \dots N-1 \\ \dfrac{dTEP_i}{dt} = a_i PCHO_i^2 + b_i PCHO_i \cdot TEP_i + t_{i-1,i} TEP_{i-1} - t_{i,i+1} TEP_i, \; i = 2, \dots N-1 \\ \dfrac{dPCHO_N}{dt} = S_N - a_N PCHO_N^2 - b_N PCHO_N \cdot TEP_N \\ \dfrac{dTEP_N}{dt} = a_N PCHO_N^2 + b_N PCHO_N \cdot TEP_N + t_{N-1,N} TEP_{N-1} - r_{sed} TEP_N \, . \end{cases} \tag{3.4}$$

The equilibria are now given by

$$TEP_1 = \frac{S_1}{t_{1,2}}, \quad TEP_i = \frac{S_i + t_{i-1,i} TEP_{i-1}}{t_{i,i+1}}, \quad TEP_N = \frac{S_N + t_{N-1,N} TEP_{N-1}}{r_{sed}} \, .$$

At some point we tried to use techniques coming from the theory of chemical reaction networks [22] to study the dynamical behavior of such models The deficiency index of these models, regarded as chemical reaction networks is equal to twice the number of depth layers -the same holds when we use the alternative kinetics (3.2). This does not allow

to apply the zero deficiency theorem, however we cannot disregard that a more thorough analysis furnishes useful applications of chemical reaction network theory to these systems. At any rate, the gross dynamical behavior can be inferred from the single layer versions, since the structure of the coupling between different layers justifies that the equilibration is transferred from the uppermost layer to the one beneath it, and sequentially in the same way until the bottom layer becomes equilibrated too.

As with the single-layer model, we ran identifiability tests for two- and three-layer models using DAISY; in both cases the models are globally identifiable. We conjecture that the same happens for an arbitrary number of layers. It would be interesting to check what implications would this fact have for the continuous limit of this model, see below.

3.3 Continuous limit

We can perform the continuous limit of (3.4) when the number of layers grows large. If the rates used in each of the layers scale with volume like mass action law rates, the limit $N \to \infty$ yields the following depth-structured model:

$$\begin{cases} \dfrac{\partial X}{\partial t} = s(x) - a(x)X^2 - b(x)X\,Y\,, \\ \\ \dfrac{\partial Y}{\partial t} = a(x)X^2 + b(x)X\,Y - \dfrac{\partial}{\partial x}[t(x)\,Y]\,. \end{cases}$$

Here the variable $x \in [0, x_{bot}]$ represents depth; these equations are complemented with the following boundary conditions after (3.4):

$$t(0) = 0, \quad t(x_{bot}) = r_{sed}\,.$$

4 Notes on nonparametric inversion for the pure coagulation equation

Here we adopt a slight modification of the strategy presented in [35]. Define the cumulative mass distribution

$$F(t,x) := \int_0^x y\,n(t,y)\,dy \quad (\partial_x F = xn(x)).$$

Tthis object satisfies the following evolution equation:

$$\frac{\partial}{\partial t}F(t,x) = \frac{1}{2}\int_0^x \left(\int_0^{x-y} \frac{\partial F}{\partial x}(t,y)\frac{\partial F}{\partial x}(t,w)K(y,w)\left(\frac{1}{y}+\frac{1}{w}\right)\,dw\right)\,dy \qquad (4.1)$$

$$-\int_0^x \left(\int_0^{\infty} \frac{\partial F}{\partial x}(t,y)\frac{\partial F}{\partial x}(t,w)\frac{K(y,w)}{w}\,dw\right)\,dy\,.$$

Introducing the profile

$$f(z) := \int_0^z wg(w)\,dw\,, \quad (\text{thus } f'(z) = z\,g(z))$$

(where z will stand for $t^\nu x$) we have

$$F(t,x) = f(t^\nu x) := f(z).$$

Replacing the profile f into (4.1) and using the homogeneity of the kernel we get to

$$\nu f'(t^\nu x)t^{\nu-1}x = \frac{1}{2}t^{(1-\lambda)\nu}\int_0^{t^\nu x}\left(\int_0^{t^\nu x-\bar{y}}\kappa(\bar{y}/\bar{w})f'(\bar{y})f'(\bar{w})\left(\frac{1}{\bar{y}}+\frac{1}{\bar{w}}\right)(\bar{w})^\lambda\,d\bar{w}\right)d\bar{y}$$

$$-t^{(1-\lambda)\nu}\int_0^{t^\nu x}\left(\int_0^{\infty}\kappa(\bar{y}/\bar{w})f'(\bar{y})f'(\bar{w})(\bar{w})^{\lambda-1}\,d\bar{w}\right)d\bar{y}\,.$$

Here the constraint

$$(1-\lambda)\nu+1=0$$

arises when demanding that the equation be independent of time. After some rearrangements, the former equation can be written in terms of $z := t^\nu x$ and $g(z)$ as follows:

$$\nu z^2 g(z) = -\int_0^1 \kappa(\xi)\left(\xi\int_{\frac{z}{1+\xi}}^{z/\xi} y^{\lambda+2}g(y)g(y\xi)\,dy + \int_{\frac{z}{1+\xi}}^{z} y^{\lambda+2}g(y)g(y\xi)\,dy\right)d\xi\,. \tag{4.2}$$

We assume that $g(z)$ and λ are known and we try to infer κ. This is a problem that fits into the following general formulation

$$data(z) = \int_0^1 \kappa(\xi)\Omega(\xi,z)\,d\xi\,. \tag{4.3}$$

We handle this problem via collocation with Tikhonov regularization. We implement the following computational procedure [10]:

1. We take a set of $M = 100$ collocation points z_i, logarithmically spaced in the range $[2^{-8}, 2^2]$. Now $data(z)$ is discretized over these points to obtain a vector $d \in \mathbb{R}^M$.

2. $\kappa(\xi)$ is expanded over a base of p B-splines $\beta_i(\xi)$ in $[0,1]$; here we use 12 cubic splines. Now we seek the expansion coefficients, which we denote $c \in \mathbb{R}^p$.

3. We compute the matrix $X_{M\times p}$, where $x_{ij} = \int_0^1 \beta_j(\xi)\Omega(\xi,z_i)\,d\xi$. Then we intend to solve the overdetermined system $X\,c^T = d$.

4. We phrase the problem as $\min_c \|Xc^T - d\|_2$ and we introduce some regularization. For that aim we consider

$$\min_c \|Xc^T - d\|_2^2 + \lambda^2 cWc^T\,.$$

   where λ^2 is the regularizing parameter and $W_{p\times p}$ is the regularization matrix. Following [35] we consider W to be composed of the scalar product of the basis elements, plus that of their first derivatives, plus that of their second derivatives -simpler choices for W do not seem to work well.

5. We seek the optimal value of λ^2 via the L-curve criterion -the generalized cross-validation criterion does not yield good results here. This we do via reduction to standard form, where instead of W we have the identity matrix. Such passage can be achieved by performing a Cholesky decomposition $W = LL^T$; then the problem in standard form can be solved in terms of the singular value decomposition $X(L^T)^{-1} = UDV^T$. We get the optimal regularizing parameter via the L-curve for the problem in standard form, solve it for the optimal λ^2 and then transform back to the original form.

4.1 Inversion in an oceanographical context

We adapt the strategy of the previous section to (2.2). As before, we set

$$F(t,x) := \int_0^x y\, n(t,y)\, dy, \quad \text{such that } \partial_x F = xn(x).$$

The evolution equation for this object reads

$$\frac{\partial F}{\partial t} = \int_0^x y \cdot \text{coagulation terms}\, dy + \int_0^x yS(y, n(t,y))\, dy\,.$$

We expect to observe some stationary F_∞. Taking the limit $t \to \infty$ in the former equation, we would obtain

$$0 = \frac{1}{2}\int_0^x \int_0^{x-y} K(y,w) yw \left(\frac{1}{y} + \frac{1}{w}\right) n_\infty(y) n_\infty(w)\, dw\, dy$$

$$- \int_0^x \int_0^\infty yK(y,w) n_\infty(y) n_\infty(w)\, dw\, dy + \text{data}(x)$$

as we have done before. Here

$$\text{data}(x) := \int_0^x yS(y, n_\infty(y))\, dy\,, \quad (\text{usually } = \mu \int_0^{\min\{x, x_{min}\}} y\, dy - \int_0^x yw(y) n_\infty(y)\, dy)$$

which is not going to be transformed any further. When we get to this point, we expect to have measured n_∞. Moreover, we assume that we have some plausible values for μ and the settling speed $w(x)$. Again, we assume that the kernel is homogeneous of degree λ. Now the homogeneity exponent is not tied to n_∞ but we will assume it to be known anyhow. We can handle the former equation to get a problem of the form

$$-data(x) = \int_0^1 \kappa(\xi) \Omega(\xi, x; \lambda, n_\infty)\, d\xi\,.$$

Here

$$\Omega(\xi, x; \lambda, n_\infty) = -\xi \int_{\frac{x}{1+\xi}}^{x/\xi} y^{\lambda+2} n_\infty(y) n_\infty(y\xi)\, dy - \int_{\frac{x}{1+\xi}}^{x} y^{\lambda+2} n_\infty(y) n_\infty(y\xi)\, dy$$

in a similar fashion to the pure coagulation problem. Thus, the computational procedure to be followed is also very similar.

5 The Lifshitz–Slyozov equation: some generalizations and potential applications in Aquatic Sciences

Here we propose a different approach to describe aggregation dynamics in aquatic ecosystems. For that aim we introduce a variant of the classical Lifshitz–Slyozov system [31], suitable to describe aggregate growth via monomer addition. This point of view seems particularly well suited to model the formation of the sea-surface microlayer (see e.g. [46]); other applications could be considered as well.

The Lifshitz–Slyozov system describes the temporal evolution of a mixture of monomers and aggregates, where individual monomers can attach to or detach from already existing aggregates. The aggregate distribution follows a transport equation with respect to a size variable, whose transport rates are coupled to the dynamics of monomers through a mass conservation relation. The initial value problem for the Lifshitz–Slyozov model thus reads

$$\begin{cases} \dfrac{\partial f(t,x)}{\partial t} + \dfrac{\partial[(a(x)u(t) - b(x))f(t,x)]}{\partial x} = 0\,, \quad t > 0\,,\ x \in (0,\infty)\,, \\ u(t) + \int_0^\infty x f(t,x)\,dx = \rho\,, \quad t > 0 \end{cases} \tag{5.1}$$

for some given $\rho > 0$, subject to the initial condition

$$f(0,x) = f^{in}(x)\,, \qquad x \in (0,\infty). \tag{5.2}$$

Here $f(t,x)$ is a nonnegative distribution of aggregates according to their size x and time t, $u(t)$ is the monomer concentration and ρ is interpreted as the total mass of the system. The kinetic rates $a(x)$ and $b(x)$ determine how fast do attachment (a given monomer attaches to a given aggregate) and detachment (a monomer detaches from a given aggregate) reactions take place. Aggregates change their size over time according to the quantity of monomers that they gain or lose through the previous reactions. Note that the attachment process is a second order kinetics, responsible of the nonlinearity, whereas detachment is a first order kinetics, as reflected in the transport term in (5.1).

The Lifshitz–Slyozov model has been traditionally used to describe late stages of phase transitions, where Ostwald ripening phenomena take place: large aggregates grow larger at the expense of smaller ones. Indeed, the classical Lifshitz–Slyozov rates are given by $a(x) = x^{1/3}$ and $b(x) = 1$; in such a case no boundary condition is needed, see e.g. [37].

However, recent applications of this framework in biologically-oriented contexts utilize a different set of kinetic rates and then a boundary condition becomes mandatory whenever the transport field points inwards. A growing literature can be found on applications to protein polymerization phenomena and neurodegenerative diseases, starting from the so-called prion model and some of its variants (see e.g. [11] and references therein), whose different versions come as modifications of the standard Lifshitz–Slyozov equations. We consider here a form of the boundary condition that describes nucleation events:

$$\lim_{x\to 0^+} (a(x)u(t) - b(x))f(t,x) = \mathfrak{n}(u(t))\,, \quad t > 0 \tag{5.3}$$

whenever $u(t) > \lim_{x\to 0^+} \frac{b(x)}{a(x)}$. The function $\mathfrak{n}$ in (5.3) can be interpreted as a nucleation rate (i.e. the rate of formation of zero-size aggregates from monomers). This rate governs the total number of aggregates as $\frac{d}{dt}\int_0^\infty f(t,x)dx = \mathfrak{n}(u(t))$ whenever $u(t) > \lim_{x\to 0^+} \frac{b(x)}{a(x)}$.

We claim in this report that this framework can be also applied to modeling in Oceanography. The sea-surface microlayer is rich in conglomerates that grow in size by an aggregation process whereby particulate organic carbon attaches to transparent exopolymeric particles; detachment effects can also take place and eventually additional terms may be included in (5.1)–(5.3), e.g. coagulation integrals. In any case, the main point is that the boundary condition (5.3) can be interpreted as the synthesis of new aggregates from monomers and not necessarily by means of a mass action kinetics. A detailed knowledge of the properties of this model (which is currently lacking) will be necessary before we can pursue our main goal: to devise inference schemes for $a(x)$, $b(x)$ and $\mathfrak{n}$.

5.1 Existence theories, long-time dynamics and rate inversion

The mathematical literature for the classical Lifshitz–Slyozov model (that is, without boundary effects) is well established, see e.g. [14, 15, 29, 30, 36, 37, 38, 42]. The works covering mathematical aspects of the initial-boundary value problem for the Lifshitz–Slyozov model are more recent. The most encompassing contribution so far is [12], arising in the framework of this project. It is important to point out that this reference sets a local-in-time theory; examples are provided that show solutions that can change from regimes where no boundary conditions are needed to descriptions where boundary effects play a role, and vice versa. Needless to say, these effects have to be taken into account before any serious undertaking of the asymptotic behavior is attempted; we are currently generalizing the mathematical framework in order to pose a global existence theory, a theory that can cope with these changes in the nucleation behavior.

The long time behavior for the classical Lifshitz–Slyozov model is analyzed in [14, 16]. Note however that our understanding of the dynamical behavior is not complete yet even in this case without boundary effects. Concerning the model with boundary conditions, it was already shown in [11] that in some particular cases the model leads to dust formation (concentration at zero size). Recently [13] we have analyzed several representative cases beyond what was done in [11], to find out that in most of them a similar concentration behavior ensues. This makes no sense from the point of view of the inference of a, b and $\mathfrak{n}$, since in experimental situations we expect to observe nontrivial steady states. It was noticed in [11] that this concentration behavior that can be somewhat prevented if fragmentation terms are incorporated into the model, which makes full sense from the oceanographic perspective. We are currently investigating this idea in more detail as a previous step to put the inverse problem on firm grounds.

6 Some notes on aggregation studies in lakes and reservoirs

We performed aggregation experiments on several samples coming from different lakes and reservoirs. Our experience during this project has shown us that the subsequent evolution of particle size as a function of time can be fit in a very reasonable way via the fitting function

$$\text{size}(t) = c_1(1 + c_2\, t)^{1/3}\,.$$

This model can be deduced from the coagulation equation using a constant coagulation kernel. Therefore, the most parsimonious description we can get from here is already accurate enough [40, 33]. Remarkably enough, the aggregation data obtained from the Cubillas reservoir (municipalities of Albolote and Atarfe, province of Granada, Spain) are fit with R^2 values circa 0.99.

During the project we measured a series of physical, chemical and biological parameters at the Cubillas reservoir, which has an important phytoplanktonic and bacterial activity. Our measurements show that acid polysaccharides (APS) act as drivers of carbon sedimentation [32]. APS concentrations showed a seasonal pattern, with maximum values associated to phytoplanktonic production. APS fluxes were mainly affected by the salinity and the total concentration of salt cations, which stimulated gel assembly, thus promoting carbon export to the bottom of the reservoir. These results are promising and can help us understand the dynamics of the global carbon cycle and its implications in the ongoing global warming.

References

[1] A. S. Ackleh, K. Deng, X. Wang, Existence-uniqueness and monotone approximation for a phytoplankton-zooplankton aggregation model. *Z. Angew. Math. Phys.* 57 (2006), 733-749

[2] H. Amann, Coagulation-fragmentation processes. *Arch. Ration. Mech. Anal.* 151 (2000), 339-366.

[3] J. Banasiak, W. Lamb, Analytic fragmentation semigroups and continuous coagulation-fragmentation equations with unbounded rates. *J. Math. Anal. Appl.* 391 (2012), 312-322.

[4] G. Bellu, M. P. Saccomani, S. Audoly, L. D'Angiò, DAISY: a new software tool to test global identifiability of biological and physiological systems. *Computer methods and programs in Biomedicine* 88 (2007), 52-61.

[5] R. Benner, R.M.W. Amon, Continuum of Major Bioelements in the Ocean. *Annu. Rev. Mar. Sci.* 7 (2015), 185-205.

[6] A. B. Burd, Modeling particle aggregation using size class and size spectrum approaches. *Journal of Geophysical Research: Oceans* 118 (2013), 3431–3443.

[7] A. B. Burd, G. A. Jackson, Modeling steady-state particle size spectra. *Environ. Sci. Tecnol.* 36 (2002), 323–327.

[8] A. B. Burd, G. A. Jackson, Particle aggregation. *Annu. Rev. Mar. Sci.* 1 (2009), 65–90.

[9] A. B. Burd, S. B. Moran, G. A. Jackson, A coupled adsorption-aggregation model of the POC/^{234}Th ratio of marine particles. *Deep-Sea Research* 47 (2000), 103-120.

[10] J. Calvo, Github: `https://github.com/JuannnCalvo`

[11] J. Calvo, M. Doumic, B. Perthame, Long-time asymptotics for polymerization models. *Comm. Math. Phys.* 363 (2018), 111–137.

[12] J.Calvo, E. Hingant, R. Yvinec, The initial-boundary value problem for the Lifshitz-Slyozov equation with non-smooth rates at the boundary. *Nonlinearity* 34 (2021), 1975-2017.

[13] J.Calvo, E. Hingant, R. Yvinec, Concentration phenomena in the Lifshitz-Slyozov model with nucleation effects. In preparation.

[14] J.-F. Collet, T. Goudon, On solutions of the Lifshitz-Slyozov model. *Nonlinearity* 13 (2000), 1239–1262.

[15] J.-F. Collet, T. Goudon, F. Poupaud, A. Vasseur, The Beker-Döring system and its Lifshitz-Slyozov limit. *SIAM J. Appl. Math.* 62 (2002), 1488–1500.

[16] J.-F. Collet, T. Goudon, A. Vasseur, Some remarks on large-time asymptotic of the Lifshitz-Slyozov equations. *J. Statist. Phys.* 108 (2002), 341–359.

[17] S.C. Doney *et al.*, Climate Change Impacts on Marine Ecosystems. *Annu. Rev. Mar. Sci.* 4 (2012), 11-37.

[18] P. B. Dubovskii, An iterative method for solving the coagulation equation with spatially inhomogeneous velocity fields. *USSR Comput. Maths. Math. Phys.* 30 (1990), 116-117.

[19] A. Engel, S. Thoms, U. Riebesell, E. Rochelle-Newall, I. Zondervan, Polysaccharide aggregation as a potential sink of marine dissolved organic carbon. *Nature* 428 (2004), 929–932.

[20] M. Escobedo, S. Mischler, B. Perthame, Gelation in coagulation and fragmentation models. *Commun. Math. Phys.* 231 (2002), 157-188.

[21] M. Escobedo, S. Mischler, M. Rodriguez Ricard, On self-similarity and stationary problem for fragmentation and coagulation models. *Annales de L'Institut Henri Poincaré - Analyse Non Linéaire* 22 (2005), 99-125.

[22] M. Feinberg, Foundations of Chemical Reaction Network Theory. Springer 2019.

[23] S. E. Flores, Modelado y Análisis Cualitativo del Proceso de Agregación en Ciencias Acuáticas. Trabajo fin de máster 2020. Universidad de Granada.

[24] N. Fournier, P. Laurençot, Existence of self-similar solutions to Smoluchowski's coagulation equation. *Commun. Math. Phys.* 256 (2005), 589-609.

[25] S. S. Fructuoso, Soluciones numéricas de la ecuación de Smoluchowski: comparativa de dos métodos utilizados y una aplicación en Oceanografía. Trabajo fin de máster 2021. Universidad de Granada. Github: `https://github.com/ssfructuoso/TFM_Ec_Coagulacion`

[26] G. A. Jackson, A. B. Burd, Aggregation in the marine environment. *Environ. Sci. Technol.*, 32 (1998), 2805–2814.

[27] E. C. Laurenceau-Cornec *et al.*, New guidelines for the application of Stokes' models to the sinking velocity of marine aggregates. *Limnology and Oceanography* 9999 (2019), 1-22.

[28] P. Laurençot, Stationary solutions to coagulation-fragmentation equations. *Annales de L'Institut Henri Poincaré - Analyse Non Linéaire* 36 (2019), 1903-1939.

[29] P. Laurençot, Weak solutions to the Lifshitz-Slyozov-Wagner equation. *Indiana Univ. Math. J.* 50 (2001), 1319–1346.

[30] P. Laurençot, S. Mischler, From the Becker-Döring to the Lifshitz-Slyozov-Wagner equations. *J. Statist. Phys.* 106 (2002), 957–991.

[31] I. M. Lifshitz, V. V. Slyozov, The kinetics of precipitation from supersaturated solid solutions. *J. Phys. Chem. Solids* 19 (1961), 35–50.

[32] A. Martínez, Acid polysaccharides as drivers of carbon sedimentation in a eutrophic reservoir. Trabajo fin de máster 2021. Universidad de Granada.

[33] A. Martínez, G. Batanero, J. Calvo, I. Reche, The role of organic matter in aggregation-sedimentation dynamics of lakes and reservoirs. A case study: Cubillas' reservoir. In preparation.

[34] I.N. McCave, i) Vertical flux of particles in the ocean. *Deep-Sea Res.* 22 (1975), 491-502, ii) Size spectra and aggregation of suspended particles in the deep ocean. *Deep-Sea Res.* 31 (1984), 329-352.

[35] R. Muralidar, D. Ramkrishna, An inverse problem in agglomeration kinetics. *Journal of Colloidal and Interface Science* 112 (1986), 348-361.

[36] B. Niethammer, On the evolution of large clusters in the Becker-Döring model. *J. Nonlinear Sci.*, 13(1):115–155, 2003.

[37] B. Niethammer, R. L. Pego, On the initial-value problem in the Lifshitz-Slyozov-Wagner theory of Ostwald ripening. *SIAM J. Math. Anal.*, 31(3):467–485, 2000.

[38] B. Niethammer, R. L. Pego, Well-posedness for measure transport in a family of nonlocal domain coarsening models. *Indiana Univ. Math. J.*, 54(2):499–530, 2005.

[39] S. C. Oukouomi Noutchie, F. E. Doungmo Goufo, Analysis of the effects of large scale marine iron fertilization. *Journal of Pure and Applied Mathematics: Advances and Applications* 8 (2012), 25-40.

[40] I. Reche, I. Peralta-Maraver, J. Calvo, I. P. Mazuecos, A. Fernández-Barbero, M. L. Pace, P. Verdugo. Calcium and iron as promoters of the self-assembly of dissolved organic matter into particles. Preprint.

[41] C.L. Sabine, T. Tanhua, Estimation of Anthropogenic CO_2 Inventories in the Ocean. *Annu. Rev. Mar. Sci.* 2 (2010), 175-198.

[42] A. Schlichting, Macroscopic limit of the Becker-Döring equation *via* gradient flows. *ESAIM Control Optim. Calc. Var.*, 25:Art. 22, 36, 2019.

[43] S. Throm, Tail behaviour of self-similar profiles with infinite mass for Smoluchowski's coagulation equation. *J. Stat. Phys.* 170 (2018), 1215–1241.

[44] M. O. Vásquez, Ecuaciones en derivadas parciales para el análisis de modelos biopoliméricos. Tesis doctoral, 2015. Universidad de Granada.

[45] P. Verdugo, Marine microgels. *Annu. Rev. Mar. Sci.* 4 (2012), 375-400.

[46] O. Wurl, L. Wurl, L. Miller, K. Johnson, and S. Vagle, Formation and global distribution of sea-surface microlayers. *Biogeosciences*, 8(1):121–135, 2011.