On the Positivity of Solutions to the Smoluchowski Equations[1]

F. P. da Costa

Instituto Superior Técnico,
Departamento de Matemática,
Av. Rovisco Pais, P-1049-001 Lisboa, Portugal
email: fcosta@math.ist.utl.pt

[1]Published in *Mathematatika,* **42**, 406-412 (1995)

1 Introduction

The dynamics of cluster growth can be modeled by the following infinite system of ordinary differential equations, first proposed by Smoluchowski, [8],

$$\dot{c}_j(t) = \tfrac{1}{2}\sum_{k=1}^{j-1} a_{j-k,k}c_{j-k}(t)c_k(t) - c_j(t)\sum_{k=1}^{\infty} a_{j,k}c_k(t), \quad j = 1, 2, \ldots, \tag{1}$$

where $c_j = c_j(t)$ represents the physical concentration of j-clusters (aggregates of j identical particles), $a_{j,k} = a_{k,j} \geq 0$ are the time-independent coagulation coefficients, measuring the effectiveness of the coagulation process between a j-cluster and a k-cluster, and the first sum in the right-hand side of (1) is defined to be zero if $j = 1$.

The quantity $\rho(c(t)) = \sum_{j=1}^{\infty} jc_j(t)$ is physically interpreted as the total density of the system modeled by (1). This suggests the following Banach space X for the study of solutions to the initial value problem for (1):

$$X = \left\{ c = (c_j) : \|c\| \stackrel{\text{def}}{=} \sum_{j=1}^{\infty} j|c_j| < \infty \right\}.$$

Moreover, since $c_j(t)$ represents a physical concentration, it is of special importance the study of non-negative solutions to (1), i.e., solutions that lie in the nonegative cone of X, $X^+ = \{c \in X : c_j \geq 0\}$.

In recent years equation (1) have attracted a good deal of mathematical interest and questions of existence, uniqueness, and asymptotic behaviour have been elucidated [1, 3, 4, 5, 6, 7]. In some of these studies, a solution to (1) in $[0, T)$ is defined to be a solution to the corresponding integral equations, [1, 3, 4],

$$c_j(t) = c_j(0) + \tfrac{1}{2}\int_0^t \sum_{k=1}^{j-1} a_{j-k,k}c_{j-k}(s)c_k(s)ds - \int_0^t c_j(s)\sum_{k=1}^{\infty} a_{j,k}c_k(s)ds, \quad j = 1, 2, \ldots$$

satisfying, for each $j \geq 1$, $c_j \in \mathcal{C}([0, T), \mathbb{R}_0^+)$, $\sum_{k=1}^{\infty} a_{j,k}c_k(\cdot) \in L^1_{\text{loc}}([0, T))$, and $\sup_{t\in[0,T)} \|c(t)\| < \infty$. We shall assume this definition in the present study.

A natural problem that arises when studying solutions to (1) is their positivity behaviour: Let $c(t) = (c_j(t))$ be a solution of (1) in $[0, T)$ with initial data $c(0) = c_0 \equiv (c_{0\,j})$. For $t \geq 0$ consider the set $\mathcal{J}(t) = \{j \in \mathbb{N}_1 : c_j(t) > 0\}$. The information on the positivity behaviour of solutions is encoded in the set-valued function $\mathcal{J}(\cdot)$, and it is clearly of some interest to be able to characterize this function, not only its t-dependence but also the way it changes with the parameters of the system and the initial data.

Clearly $\mathcal{J}(\cdot)$ will depend on the coefficients $a_{j,k}$ since if, for instance, there is a constant $N \geq 1$ such that $a_{j,k} = 0$ for $j + k > N$, then (1) is effectively reduced to an N-dimensional system of ordinary differential equations for the phase variables $(c_1, \dots, c_N)$ with the remaining variables satisfying the trivial system $\dot{c}_j \equiv 0$, $j \geq N + 1$, and thus, for all $t > 0$, $\#\mathcal{J}(t) \leq N$ if $c_{0\,j} = 0$ for $j \geq N + 1$. Thus, by changing the size N of the *truncation* we clearly can obtain sets $\mathcal{J}(t)$ of different sizes. We will return to these truncated cases in the end of the paper, but until then we shall consider (1) to be a genuinely infinite dimensional system, and we assume the positivity condition $a_{j,k} > 0$ for all j and k, which is the most interesting case from the point of view of applications.

We prove that, for all $t > 0$, $\mathcal{J}(t)$ is constant, $\mathcal{J}(t) \equiv \mathcal{J}$, and the set $\mathcal{J}$ is infinite and independent of the parameters $a_{j,k}$, provided the positivity assumption holds. This implies, physically, an infinite velocity of reaction between different clusters, and in this way highlights one of the limitations of the mean field model (1).

The proof actually gives a complete characterization of the set $\mathcal{J}$ in terms of the positivity properties of the initial data $\mathcal{J}(0)$, namely we prove that $j \in \mathcal{J}$ if and only if there exists elements $p_1, \dots, p_n \in \mathcal{J}(0)$, $n < j$, and positive integers $m_1, \dots, m_n$ such that $j = \sum_{i=1}^{n} m_i p_i$.

We shall use the following notation: let $c(t)$ be a solution of (1) in $[0, T)$ with initial data $c(0) = c_0$, define

$$\mathcal{J}(t) = \{j \in \mathbb{N}_1 : c_j(t) > 0\}$$

$$P = \mathcal{J}(0)$$

$$P_m = P \bigcap \{1, \dots, m\}$$

$$\operatorname{span}_{\mathbb{N}_0}(P) = \left\{ j = \sum_i n_i p_i : p_i \in P,\ n_i \in \mathbb{N}_0, \text{ and } \max_i n_i > 0 \right\}$$

$$\mathbb{N}_1 \cdot P = \{ j = np : n \in \mathbb{N}_1, p \in P \}$$

2 Results

The main result is the following

Theorem 1

Assume $a_{j,k} > 0$ *for all* j *and* k. *Let* $T \in (0, \infty]$, *and let* c *be a solution of* (1) *on* $[0, T)$ *with initial data* $c(0) = c_0 \neq 0,\ c_0 \in X^+$.

Then, for all $t > 0$, $\mathcal{J} \equiv \mathcal{J}(t)$ *is independent of* t, *and* $\mathcal{J} = \operatorname{span}_{\mathbb{N}_0}(P)$.

We prove Theorem 1 in several lemmas. In the first two lemmas we obtain that, for each $t > 0$, the set $\mathcal{J}(t)$ is infinite and $\mathcal{J}(t) \supseteq \mathcal{J}(\tau)$ if $t \geq \tau \geq 0$. Then, in a series of lemmas, we establish that, for all $t \in (0, T)$, the set $\mathcal{J}(t)$ is given by $\operatorname{span}_{\mathbb{N}_0}(P)$ and thus is independent of t.

Lemma 1 (Theorem 2.1 of [3])

With the assumptions of Theorem 1 we have that for all $t \in (0, T)$ *there exists a positive integer sequence* $j_k(t) \underset{k \to \infty}{\longrightarrow} \infty$ *such that* $c_{j_k(t)}(t) > 0$ *for all positive integers* k.

Proof: Assume the contrary, i.e., there exists a $\tau > 0$ and an integer $L = L(\tau) > 0$ such that $c_j(\tau) = 0$ for all $j \geq L$. Let

$$\phi_j(t) = \sum_{k=1}^{\infty} a_{j,k}\, c_k(t), \quad R_j(t) = \tfrac{1}{2} \sum_{k=1}^{j-1} a_{j-k,k}\, c_{j-k}(t)\, c_k(t),$$

and

$$E_j(t) = \exp(\int_0^t \phi_j(s)\, ds).$$

Thus, equation (1) for $j = L$ can be written as

$$0 = c_L(\tau)E_L(\tau) = c_{0L} + \int_0^\tau E_L(s)R_L(s)\, ds.$$

As $\phi_j(t), R_j(t) \geq 0$ and $E_j(t) \geq 1$ for all t, this implies $c_{0L} = 0$ and $R_L(s) = 0$ a.e. $s \in [0, \tau]$. In fact, $R_L \equiv 0$ on $[0, \tau]$ since R_L is a finite sum of continuous functions, and this implies $c_j \equiv 0$ on $[0, \tau]$ for all $j < L$; for suppose there is a J such that $\lfloor \frac{L+1}{2} \rfloor \leq J < L$, and c_J is not identically zero on $[0, \tau]$, where $\lfloor x \rfloor$ denotes the integer part of x. Since $2J \geq L+1$, repeating the argument above for $j = 2J$ we obtain $R_{2J} \equiv 0$ on $[0, \tau]$ and, since R_{2J} contains a term $c_J c_J$, c_J must be identically zero. Thus, we have $c_j \equiv 0$ for all $j \geq \lfloor \frac{L+1}{2} \rfloor$. Now applying the same argument with L substituted by $\lfloor \frac{L+1}{2} \rfloor$ and proceeding as above we obtain $c_j(t) = 0$ for all $t \in [0, \tau], j = 1, 2, \cdots$. But then $c_0 = (c_j(0)) = 0$, contrary to the assumption. This proves the result. ■

Remark 1 Fix $t > 0$ and let $\{j_k(t)\}$ denote the set of elements of the sequence given in Lemma 1. Considering all possible such sequences with the usual set inclusion relation, we can define a maximal sequence $\mathcal{J}(t)$ satisfying the Lemma. From now on we shall always assume the sequences under consideration to be maximal for each $t \in (0, T)$; also, in order not to overload the notation, we shall write $t > 0$ instead of $t \in (0, T)$.

Lemma 2

With the assumptions of Theorem 1 we have $\mathcal{J}(t) \supseteq \mathcal{J}(\tau)$, for all $t \geq \tau \geq 0$.

Proof: Let c be a solution of (1) on $[0, T)$ and suppose that, for a given $\tau \geq 0$ and positive integer j we have $j \in \mathcal{J}(\tau)$. Then, for all $t > \tau$,

$$c_j(t)E_j(t) = c_j(\tau)E_j(\tau) + \int_\tau^t E_j(s)R_j(s)ds \geq c_j(\tau)E_j(\tau) > 0,$$

and hence $j \in \mathcal{J}(t)$ for all $t > \tau$. ■

Lemma 3

With the assumptions of Theorem 1, *for every* $t > 0$, $\mathcal{J}(t) = \mathbb{N}_1$ *if and only if* $1 \in P$.

Proof: Suppose $1 \notin P$. Then, for all $t > 0$, $c_1(t)E_1(t) = c_1(0) = 0$, which implies $c_1 \equiv 0$ on $[0, T)$, i.e., $1 \notin \mathcal{J}(t)$, and thus $\mathcal{J}(t) \neq \mathbb{N}_1$.
Conversely, let $j \notin \mathcal{J}(t)$ for some $t > 0$. Then,

$$0 = c_j(t)E_j(t) = c_j(0) + \int_0^t E_j(s)R_j(s)ds, \tag{2}$$

and, by continuity of $R_j(\cdot)$, this implies $R_j \equiv 0$ on $[0, t]$. Since solutions are non-negative we must have, for all $k \in \{1, \dots, j-1\}$

$$c_{j-k}(s)c_k(s) = 0 \quad \text{for all} \quad s \in [0, t].$$

In particular $c_{j-1}c_1$ must be identically zero on $[0, t]$. Now, suppose $c_1(0) > 0$. Then, by Lemma 2, $c_1(s) > 0$ for all $s \geq 0$ and thus c_{j-1} must be identically zero on $[0, t]$. Write (2) with j substituted by $j - 1$. The same argument can be applied repeatedly and we finally conclude that all $c_{j-1}, c_{j-2}, \dots, c_2$ are identically zero on $[0, t]$. But then, for all $s \in [0, t]$,

$$0 = c_2(s)E_2(s) = \frac{1}{2}\int_0^s E_2(\varsigma)(c_1(\varsigma))^2 d\varsigma,$$

which implies $c_1 \equiv 0$ on $(0, t]$, and by continuity $c_1(0) = 0$. This contradiction proves the Lemma. ■

Lemma 4

With the assumptions of Theorem 1, $\mathbb{N}_1 \cdot P \subseteq \mathcal{J}(t)$, *for all* $t > 0$.

Proof: Let $j = np$ for some $n \in \mathbb{N}_1$ and $p \in P$.

(i) If $n = 1$ we have Lemma 2.

(ii) Suppose $n = 2$. Then $j = 2p$ and for all $t > 0$,

$$\begin{aligned} c_j(t)E_j(t) &= c_j(0) + \int_0^t E_j(s)R_j(s)ds \geq \\ &\geq \int_0^t E_j(s)R_j(s)ds \geq \\ &\geq \frac{1}{2}\int_0^t E_j(s)a_{j-p,p}c_{j-p}(s)c_p(s)ds = \\ &= \frac{1}{2}\int_0^t E_j(s)a_{p,p}(c_p(s))^2 ds > 0, \end{aligned}$$

and so $j \in \mathcal{J}(t)$.

(iii) Now proceed by induction: let $n \geq 3$ and assume that $(n-1)p \in \mathcal{J}(t)$ for all $t > 0$. Then

$$\begin{aligned} c_j(t)E_j(t) &= c_j(0) + \int_0^t E_j(s)R_j(s)ds \geq \\ &\geq \int_0^t E_j(s)R_j(s)ds \geq \\ &\geq \int_0^t E_j(s)a_{j-p,p}c_{j-p}(s)c_p(s)ds = \\ &= \int_0^t E_j(s)a_{(n-1)p,p}c_{(n-1)p}(s)c_p(s)ds > 0 \end{aligned}$$

and so $j \in \mathcal{J}(t)$. ■

Lemma 5

With the assumptions of Theorem 1, $\mathrm{span}_{\mathbb{N}_0}(P) \subseteq \mathcal{J}(t)$, *for all* $t > 0$.

Proof: Let $j_0 \in \mathrm{span}_{\mathbb{N}_0}(P)$. Clearly, since the coefficients in the expansion of j_0 are non-negative, we have in fact $j_0 \in \mathrm{span}_{\mathbb{N}_0}(P_{j_0})$, i.e., there exist positive integers n_i and elements of P_{j_0}, p_i, such that

$$j_0 = \sum_{i=1}^{L} n_i p_i \tag{3}$$

where $L \leq j_0$ is an integer. If $L = 1$ then $j_0 \in \mathbb{N}_1 \cdot P$ and the result follows by Lemma 4. Suppose $L \geq 2$. Define $j_k = j_{k-1} - n_k p_k$ for $k = 1, \ldots, L-1$. Then $j_k = j_0 - \sum_{i=1}^{k} n_i p_i = \sum_{i=k+1}^{L} n_i p_i$. We prove that $j_k \in \mathcal{J}(t)$ for all $k = 0, 1, \ldots, L-1$, and all $t > 0$.

(i) Let $k = L-1$. Then $j_{L-1} = n_L p_L \in \mathbb{N}_1 \cdot P$ so, by Lemma 4, $j_{L-1} \in \mathcal{J}(t)$.

(ii) Assume $j_k \in \mathcal{J}(t)$, for all $t > 0$. We prove that $j_{k-1} \in \mathcal{J}(t)$, for all $t > 0$:

$$\begin{aligned} c_{j_{k-1}}(t) E_{j_{k-1}}(t) &= c_{j_{k-1}}(0) + \int_0^t E_{j_{k-1}}(s) R_{j_{k-1}}(s) ds \geq \\ &\geq \int_0^t E_{j_{k-1}}(s) a_{j_{k-1}-n_k p_k, n_k p_k} c_{j_{k-1}-n_k p_k}(s) c_{n_k p_k}(s) ds = \\ &= \int_0^t E_{j_{k-1}}(s) a_{j_k, n_k p_k} c_{j_k}(s) c_{n_k p_k}(s) ds > 0, \end{aligned}$$

since $n_k p_k \in \mathbb{N}_1 \cdot P \subseteq \mathcal{J}(t)$. ■

Lemma 6

With the assumptions of Theorem 1, $\mathcal{J}(t) \subseteq \mathrm{span}_{\mathbb{N}_0}(P)$, for all $t > 0$.

Proof: By Lemma 4 we know that $\mathbb{N}_1 \cdot P \subseteq \mathcal{J}(t)$. Clearly $\mathbb{N}_1 \cdot P \subseteq \mathrm{span}_{\mathbb{N}_0}(P)$ so it is sufficient to prove that, if $\mathcal{J}(t) \backslash (\mathbb{N}_1 \cdot P) \neq \emptyset$, then $\mathcal{J}(t) \backslash (\mathbb{N}_1 \cdot P) \subseteq \mathrm{span}_{\mathbb{N}_0}(P)$.

Let $\mathcal{M}(t) = \mathcal{J}(t) \backslash (\mathbb{N}_1 \cdot P)$, and let $m \in \mathbb{N}_1 \cup \{\infty\}$ be the number of elements of $\mathcal{M}(t)$. Write $\mathcal{M}(t) = \{\mu_i\}$ with $\mu_1 < \mu_2 < \ldots$. We prove by induction that, for all i, $\mu_i \in \mathrm{span}_{\mathbb{N}_0}(P)$. Observe first that $1 \notin \mathcal{M}(t)$: if $1 \in \mathcal{J}(t) \backslash (\mathbb{N}_1 \cdot P)$ then $1 \notin P$ but then, by the proof of Lemma 3, $1 \notin \mathcal{J}(t)$, a contradiction.

(i) Consider μ_1. We have

$$\begin{aligned} 0 \ < \ c_{\mu_1}(t)E_{\mu_1}(t) \ &= \ c_{\mu_1}(0) + \int_0^t E_{\mu_1}(s)R_{\mu_1}(s)ds = \\ &= \ \frac{1}{2}\int_0^t E_{\mu_1}(s)\sum_{k=1}^{\mu_1-1} a_{\mu_1-k,k}c_{\mu_1-k}(s)c_k(s)ds. \end{aligned} \tag{4}$$

By Lemma 2 this implies that at least one of the terms in the sum in the right-hand side is positive on $(\tau, t]$ for some $\tau \in [0, t)$ i.e., there exists a $\tilde{k}_1 \in \{1, \ldots, \mu_1 - 1\}$ such that $c_{\tilde{k}_1}(s)c_{\tilde{k}_2}(s) > 0$, for all $s \in (\tau, t]$, where $\tilde{k}_2 = \mu_1 - \tilde{k}_1$. In particular $\tilde{k}_1$ and $\tilde{k}_2$ are in $\mathcal{J}(t)$. Since $\tilde{k}_1$, $\tilde{k}_2 < \mu_1$ and $\mu_1 = \min \mathcal{M}(t)$, we have $\tilde{k}_1$, $\tilde{k}_2 \in P_{\mu_1-1}$, and so there exists non-negative integers n_1 and n_2, and elements of P_{μ_1-1}, p_1 and p_2, such that

$$\mu_1 = \tilde{k}_1 + (\mu_1 - \tilde{k}_1) = \tilde{k}_1 + \tilde{k}_2 = n_1p_1 + n_2p_2 \in \mathrm{span}_{\mathbb{N}_0}(P_{\mu_1-1}).$$

(ii) Assume $\mu_q \in \mathrm{span}_{\mathbb{N}_0}(P)$ for all $q \in \{1, \ldots, i-1\}$. Then, by the argument in (i) with μ_1 substituted by μ_i we have $\mu_i = n_1p_1 + n_2p_2$, for some non-negative integers n_1, n_2, and some $p_1, p_2 \in P_{\mu_i-1} \cup \{\mu_1, \ldots, \mu_{i-1}\}$ and thus $\mu_i \in \mathrm{span}_{\mathbb{N}_0}(P)$.

Hence, $\mathcal{J}(t) = \mathbb{N}_1 \cdot P \cup \mathcal{M}(t) \subseteq \mathrm{span}_{\mathbb{N}_0}(P)$. ■

Proof of Theorem 1: By Lemma 1, $\mathcal{J}(t)$ is an infinite set for all $t > 0$. Moreover, for all $t > 0$, $\mathcal{J}(t) \supseteq \mathrm{span}_{\mathbb{N}_0}(P)$, (by Lemma 5) and $\mathcal{J}(t) \subseteq \mathrm{span}_{\mathbb{N}_0}(P)$ (by Lemma 6). This concludes the proof. ■

3 Final Remarks

As promised in the Introduction let us briefly return to the truncated N-dimensional system obtained from (1) by assuming

$$a_{j,k} = 0 \quad \text{if and only if} \quad j + k > N, \tag{5}$$

for some fixed $N \geq 1$, namely,

$$\begin{cases} \dot{c}_j = \frac{1}{2}\sum_{k=1}^{j-1} a_{j-k,k}c_{j-k}c_k - c_j\sum_{k=1}^{N-j} a_{j,k}c_k, & j = 1, 2, \ldots, N \\ \\ \dot{c}_j = 0, & j \geq N+1. \end{cases} \tag{6}$$

It is easy to conclude from the proofs in Section 2 that the following result holds

Theorem 2

Assume (5)*. Then, for all* $t > 0$, $\mathcal{J}(t) \equiv \mathcal{J}$ *is independent of* t*, and*

$$\mathcal{J} = P \bigcup \left(\{1, \ldots, N\} \cap \mathrm{span}_{\mathbb{N}_0}(P)\right).$$

The truncated problem (6) is but one case where the assumption $a_{j,k} > 0$ for all j and k is not satisfied. Other cases include the diagonal system ($a_{j,k} > 0$ if and only if $j = k$, [2]) and the generalized Becker-Döring equations ($a_{j,k} = 0$ if and only if $\min\{j, k\} > N$ for some constant $N \geq 1$, [4]). For these cases, using the above arguments, we can still prove that $\mathcal{J}(t)$ is an infinite set independent of t for $t > 0$, but the caracterization of $\mathcal{J}$ will differ from the one given in the present paper. For instance, in the diagonal case, if the initial data satisfies $c_j(0) > 0$ if and only if $j = m$, for some integer m, it is easy to prove that, for all $t > 0$, $\mathcal{J} = \left\{2^k m : k \in \mathbb{N}_0\right\}$. For the generalized Becker-Döring equation the caracterization of $\mathcal{J}$ in terms of P and N is already a quite difficult problem for general $N > 1$.

Acknowledgements

I thank R. Dunwell for a useful remark that led to the simplification of the original proof of Lemma 6.

The author was supported by Fundação Calouste Gulbenkian, grant 9/90/B

References

[1] Ball, J., Carr, J.: 'The discrete coagulation-fragmentation equations: existence, uniqueness, and density conservation', *J. Stat. Phys.,* **61**, 203-234 (1990).

[2] Buffet, E, Pulé, J.: 'Gelation: the diagonal case revisited', *Nonlinearity,* **2**, 373-381 (1989).

[3] Carr, J., da Costa, F.P.: 'Instantaneous gelation in coagulation dynamics', *Z. Angew. Math. Phys,* **43**, 974-983 (1992).

[4] Carr, J., da Costa, F.: 'Asymptotic behavior of solutions to the coagulation-fragmentation equations. II. Weak fragmentation', *J. Stat. Phys.,* **77**, 89-123 (1994).

[5] Leyvraz, F.: 'Existence and properties of post-gel solutions for the kinetic equations of coagulation', *J. Phys. A: Math. Gen.,* **16**, 2861-2873 (1983).

[6] Leyvraz, F., Tschudi, H.R.: 'Singularities in the kinetics of coagulation processes', *J. Phys. A: Math. Gen.,* **14**, 3389-3405 (1981).

[7] Slemrod, M.: 'A note on the kinetic equations of coagulation', *J. Integral Eq. and Appl.,* **3**, 167-173 (1991).

[8] Smoluchowski, M.: 'Drei Vorträge über Diffusion, Brownsche Bewegung und Koagulation von Kolloidteilchen', *Physik Z.,* **17**, 557-585 (1916).