The effect of stochastic correlations and fluctuations in the collision-coalescence process revisited: An algorithm for the numerical solution of the master equation and numerical results for realistic kernels.

L. Alfonso

Universidad Autónoma de la Ciudad de México, México City, 09790, México

Generally, the modeling of size distribution in a collision-coalescence system is performed by the Smoluchowski equation or kinetic collection equation, which is a deterministic equation and has no stochastic correlations or fluctuations included. However, the full stochastic description of the growth of cloud particles in a coalescing system can be obtained from the solution of the master (or V- equation), which models the evolution of the state vector for the number of droplets of a given mass. Due to its complexity, only limited results were obtained for certain type of kernels (sum, product and constant kernels). In this work, a general algorithm for the solution of the master equation for stochastic coagulation was proposed. The performance of the method was checked by comparing the time evolution for the state probabilities with the analytical results obtained by other authors. Fluctuations and correlations were calculated for the hydrodynamic kernel, and true stochastic averages obtained from the master equation were compared with numerical solutions of the kinetic collection equation for that case.

1. INTRODUCTION

The evolution of the size distribution of coalescing particles has often been described by the Smoluchowski coagulation equation, known under a number of names ("kinetic", "collection" and "coalescence"). The discrete form of this equation has the form [1]:

$$\frac{\partial N(i,t)}{\partial t} = \frac{1}{2}\sum_{j=1}^{i-1} K(i-j,j)N(i-j)N(j) - N(i)\sum_{j=1}^{\infty} K(i,j)N(j) \qquad (1)$$

where $N(i,t)$ is the average number of droplets with mass x_i, and $K(i,j)$ is the collection kernel related to the probability of coalescence of two droplets of masses x_i and x_j. In Eq. (1), the time rate of change of the average number of droplets with mass x_i is determined as the difference between two terms: the first term describes the average rate of production of droplets of mass x_i due to coalescence between pairs of drops whose masses add up to mass x_i, and the second term describes the average rate of depletion of droplets with mass x_i due to their collisions and coalescence with other droplets.

Within the Smoluchowski approach (1), it is assumed that fluctuations in the concentrations are negligible small. This assumption can only be correct if the volume and the number of particles are infinite large. An alternative approach considers the coalescence process in a system of finite number of particles, with fluctuations that are no longer negligible. This finite-volume description is intrinsically stochastic and has been pioneered by Marcus [2] Bayewitz et al. [3] and studied in detailed by Lushnikov [4, 5] and more recently by Tanaka and Nakazawa [6].

Within this approach a system of particles whose total mass is M_T is considered. The mass distribution of the particles is described by giving the number n_i of particles with mass i, i.e. $n_1, n_2, n_3, \ldots, n_N$. Then, the state of the mass distribution of the particle system is described by N dimensional state vector $\bar{n} = (n_1, n_2, ..., n_N)$ and the time evolution of the joint probability $P(n_1, n_2, ..., n_N; t)$ that the system is in state $\bar{n} = (n_1, n_2, ..., n_N)$ at time t is calculated according to the equation [5]:

$$\begin{aligned}\frac{\partial P(\bar{n})}{\partial t} &= \sum_{i=1}^{N}\sum_{j=i+1}^{N} K(i,j)(n_i+1)(n_j+1)P(...,n_i+1,...,n_j+1,...,n_{i+j}-1,...;t) \\ &+ \sum_{i=1}^{N}\frac{1}{2}K(i,i)(n_i+2)(n_i+1)P(...,n_i+2,...,n_{2i}-1,...;t) \\ &- \sum_{i=1}^{N}\sum_{j=i+1}^{N} K(i,j)n_i n_j P(\bar{n};t) - \sum_{i=1}^{N}\frac{1}{2}K(i,i)n_i(n_i-1)P(\bar{n};t) \qquad (2)\end{aligned}$$

The master or V-equation (2) is a gain-loss equation for the probability of each state $\bar{n} = (n_1, n_2, ..., n_N)$. The sum of the first two terms is the gain due to transition from other states, and the sum of the last two terms is the loss due to transitions into other states. The gain terms show that the system may be reached from any state with an i-mer and a j-mer more, and one $(i+j)$-mer less. The transition rates are $K(i,j)(n_i+1)(n_j+1)$ if $i \neq j$ and $K(i,i)(n_i+1)(n_i+2)$ if $i = j$. From conservation of the total probability, $P(\bar{n};t)$ must satisfy the relation:

$$\sum_{\bar{n}} P(\bar{n};t) = 1 \qquad (3)$$

Where the sum is taken over all states. Additionally, the total mass of the system must be conserved, and the particle number n_i is non negative for any mass i. Thus, we have:

$$\sum_{i=1}^{N} i n_i = M\ , \quad n_i \geq 0, \qquad i = 1,...,N \qquad (4)$$

The exact solution of the master equation (2) is only known for a limited number of cases (constant, sum and product kernels) and for monodisperse initial conditions. For this special cases the master equation has been solved by Lushnikov [3, 4] and Tanaka and Nakazawa [5] in terms of the generating function

of $P(\bar{n};t)$. For general, nonmonodisperse initial conditions, the solution of (2) is not known.

Additionally, for stochastic coagulation, approximate solutions were calculated by using the system size expansion or Ω-expansion [7], which permits to find solutions of (2) valid in the limit of a large system. However, this method gives less reliable results in a system with a low number of particles.

Then, in order to obtain solutions for more realistic kernels (Brownian motion, differential sedimentation etc.), it has to be solved numerically.

In this paper, we present a novel numerical algorithm to obtain the full solution of the master equation for any type of kernels and general nonmonodisperse initial conditions.

2. THE NUMERICAL ALGORITHM

To solve equation (2) by finite differences, the joint probability $P(n_1, n_2, ..., n_N; t)$ must be discretized into a multidimensional array. The main drawback of this approach is its susceptibility to the curse of dimensionality, i.e. the exponential growth in memory and computational requirements in the number of problem dimensions.

For example, for a system with a mono-disperse initial condition $P(50,0,0,...,0;0)=1$, even with the restriction (4), we would be in need to define a 50 dimensional array with about $3\text{x}10^{20}$ elements, which is computationally prohibitive.

2.1 Calculation of all possible configurations

Instead of the brute force discretization of the multi-dimensional joint probability distribution, the solution to this problem lies on the generation of all possible states from an initial configuration, and the posterior calculation of the time evolution of the probability $P(\bar{n};t)$ for each generated configuration by using the equation master equation (2). From an arbitrary initial condition $P(n_{01}, n_{02}, ..., n_{0N}; 0)=1$ all possible states can be generated numerically, knowing that the only allowed transitions are of the form $\bar{n}^{(1)} \to \bar{n}$ if $i \neq j$ and $\bar{n}^{(2)} \to \bar{n}$ $i = j$, where $\bar{n}^{(1)}$ and $\bar{n}^{(2)}$ are the state vectors:

$$\bar{n}^{(1)} = (n_1, ..., n_i + 1, ..., n_j + 1, ..., n_{i+j} - 1, ..., n_N) \quad (5a)$$

$$\bar{n}^{(2)} = (n_1, ..., n_i + 2, ..., n_{2i} - 1, ..., n_N) \quad (5b)$$

Then, for a system consisting of N monomers at $t=0$, $R(N)$ states can be realized, where $R(N)$ is the number of solutions in integers $\bar{n}$ of Eq. (4) for the conservation of mass. The total number of states can be approximated from the equation [8]:

$$R(N) \sim \frac{1}{4N\sqrt{3}} \exp\left(\pi \left(2N/3\right)^{1/2}\right) \quad (6)$$

Note that, although $R(N)$ increases very quickly with N (for example, $R(50)$=217590 and $R(100)$=190 569 232), a number of states that is perfectly manageable with an average computer is obtained (compared with the 50 dimensional array with $3\text{x}10^{20}$ elements required for N=50). The state generation algorithm stops once the terminal state $\bar{n} = (0,0,0,...,1)$ is obtained. The formula (6) although slightly overestimate the number of states gives estimates that can be used in order to check the performance of the algorithm. For N=6, 10, 20 we obtained 11, 42 and 627 configurations respectively, and 13, 48 and 692 by using (6). As an example, the 11 possible configurations generated from the initial state $(6,0,0,0,0,0)$ are displayed in Fig. 1.

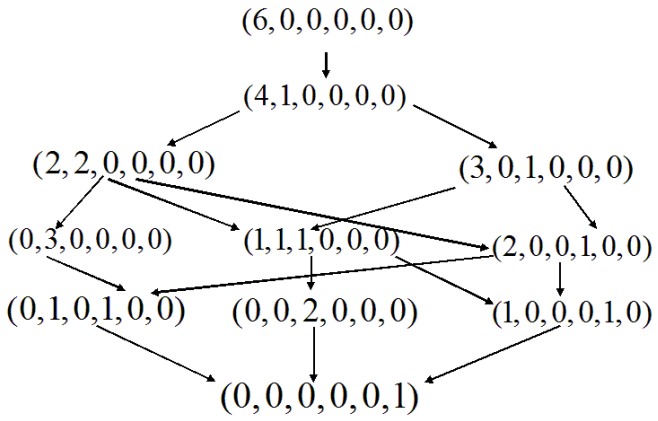

Figure 1. States space for the initial condition $P(6,0,0,0,0,0;0)=1$ with the constraint $\sum_{i=1}^{N} i n_i = 6$.

2.2 Time evolution of the probability distribution $P(\bar{n};t)$ for each state

At t_0=0, for the initial state $P(n_{01}, n_{02}, n_{03}, n_{04}, ...; t_0)=1$, and the probabilities for the rest of the states are set equal to 0. The probabilities of all previously calculated configurations are updated according to the first order finite difference scheme:

$$\begin{aligned} P\left(\bar{n}; t_0 + \Delta t\right) = P\left(\bar{n}; t_0\right) \\ + \Delta t \sum_{i=1}^{N} \sum_{j=i+1}^{N} K(i,j)(n_i + 1)(n_j + 1) \\ \times P(..., n_i + 1, ..., n_j + 1, ..., n_{i+j} - 1, ...; t_0) \\ + \Delta t \sum_{i=1}^{N} \frac{1}{2} K(i,i)(n_i + 2)(n_i + 1) \\ \times P(..., n_i + 2, ..., n_{2i} - 1, ...; t_0) \\ - \Delta t \sum_{i,j=1}^{N} K(i,j) n_i n_j P(\bar{n}; t_0) \\ - \Delta t \sum_{i=1}^{N} \frac{1}{2} K(i,i) n_i (n_i - 1) P(\bar{n}; t_0) \end{aligned} \quad (7)$$

It is clear from (7) that the state probabilities at $t=t_0+\Delta t$ will increased if the states from which transitions are allowed, have a non-zero probability at $t = t_0$ (second and third terms in the rhs of Eq. (7)), and will decrease due to collisions of particles from the same state at $t = t_0$ (fourth term and fifth terms in the rhs of Eq. (7)) if $P(\bar{n};t_0)$ is positive. The finite difference equation for $P(1,0,0,0,1,0)$ was written to illustrate the method. As can be checked from the generation scheme displayed in Figure 1, the only allowed transitions to $(1,0,0,0,1,0)$ are from the states $(1,1,1,0,0,0)$ and $(2,0,0,1,0,0)$. Then, at $t=t_0+\Delta t$, $P(0,1,0,1,0,0;t_0+\Delta t)$ will increase if $P(1,1,1,0,0,0;t_0)$ and $P(2,0,0,1,0,0;t_0)$ are non-zero at $t = t_0$. On the other hand, $P(1,0,0,0,1,0;t_0+\Delta t)$ will decrease due to collisions from particles within the same state at $t = t_0$. Then, at each time step, $P(1,0,0,0,1,0;t_0+\Delta t)$ is calculated from the finite difference equation:

$$\begin{aligned} P(1,0,0,0,0,1,0;t_0+\Delta t) = {} & P(1,0,0,0,0,1,0;t_0) \\ & + \Delta t K(2,3)(n_2+1)(n_3+1)P(1,1,1,0,0,0;t_0) \\ & + \Delta t K(1,4)(n_1+1)(n_4+1)P(2,0,0,1,0,0;t_0) \\ & - \Delta t K(1,5)(n_1+1)(n_5+1)P(1,0,0,0,1,0;t_0) \end{aligned} \quad (8)$$

As an exercise, the time evolution of each state was calculated for the coalescence kernel $K(i,j)=(i^{1/2}+j^{1/2})/40$ from Marcus paper [2]. The results obtained for six of the states are displayed in Fig. 2.

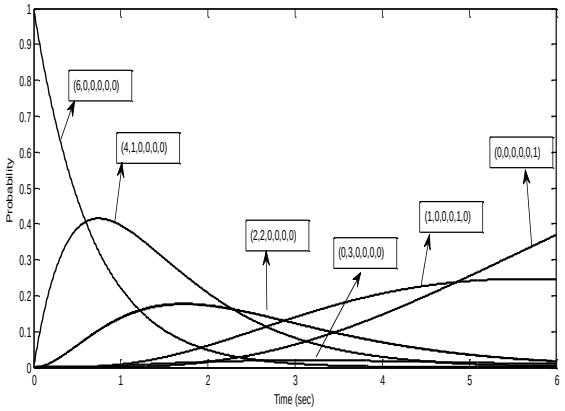

FIG. 2. Time evolution of the probability for six of the states for the initial condition $P(6,0,0,0,0,0;0)=1$.

2.3 Calculation of the expected values of the number of particles for each particle mass

The number of particles for a given mass n_1, n_2, …,n_N are discrete random variables whose probability distributions can be written as:

$$P(n,m;t) = \sum_{Except \;\; x_m} P\left(n_1,n_2,...,n_m=n,...;t\right) \quad (9)$$

Usually, the numerical implementation of (9) would involve calculating the sum of all elements of a multidimensional array, which is computationally very expensive. Our approach is simpler: Once the probabilities of all possible states are determined for all times, $P(n,m;t)$ can be calculated just by summing over all states that have $n_m=n$:

$$P(n,m;t) = \sum_{All \;\; states \;\; with \;\; n_m=n} P\left(n_1,n_2,...,n_m=n,...;t\right) \quad (10)$$

The expected value $\langle n_m \rangle$ for the number of particles of mass m is then calculated from the equation:

$$\langle n_m \rangle = \sum_n nP(n,m;t) \quad (11)$$

As an example, for the system from Fig. 1, the procedure for calculating the probability distribution $P(n,1;t)$ (probability of finding n particles of size $m=1$ at time t) is displayed in Table 3.

Table 1. Probability distribution $P(n,1;t)$ of finding n particles of size $m=1$ at time t, for a system with the initial condition $P(6,0,0,0,0,0;0)=1$.

Probability distribution $P(n,1;t)$
$P(0,1;t)=P(0,1,0,1,0,0,t)+P(0,0,2,0,0,0)+P(0,0,0,0,0,1)$
$P(1,1;t)=P(1,1,1,0,0,0,t)+P(1,0,0,0,1,0)$
$P(2,1;t)=P(2,2,0,0,0,0;t)+P(2,0,0,0,1,0;t)$
$P(3,1;t)=P(3,0,1,0,0,0;t)$
$P(4,1;t)=P(4,1,0,0,0,0;t)$
$P(5,1;t)=0$
$P(6,1;t)=P(6,0,0,0,0,0;t)$

3. COMPARISON WITH ANALYTICAL SOLUTIONS AND PRELIMINAR NUMERICAL RESULTS FOR REALISTIC KERNELS

We have tested the numerical code against the exact size distribution of the master equation reported in Tanaka and Nakazawa [6] for the sum kernel $K(i,j)=B(x_i+x_j)$, with the monodisperse initial condition $P(N_0,0,0,...,0;0)$:

$$\langle n_m \rangle = C_m^{N_0}\left(\frac{i}{N_0}\right)^{m-1}\left\{1-\frac{m}{N_0}\left(1-e^{T}\right)\right\}^{N_0-m-1}\times\left(1-e^{-T}\right)^{m-1}e^{-T} \quad (12)$$

In (12), N_0 is the initial number of particles, v_0 is the initial volume of droplets, $C_m^{N_0}$ is the binomial coefficient, $T=BN_0v_0t$, B=8.82×10^2 cm^3 sec^{-1} and $\langle n_m \rangle$ are the true

stochastic averages for each particle mass *m*. Turning to concrete numerical examples, the droplet size distribution for a system with N_0=10 droplets of 10 μm in radius (droplet mass 4.189×10^{-9}g) at *t*=0, was calculated at *t*=500 and 1000 sec. by using the numerical algorithm. The time step was set equal to Δt=0.1 sec. As can be observed in Fig. 3 an excellent agreement was founded between the two approaches.

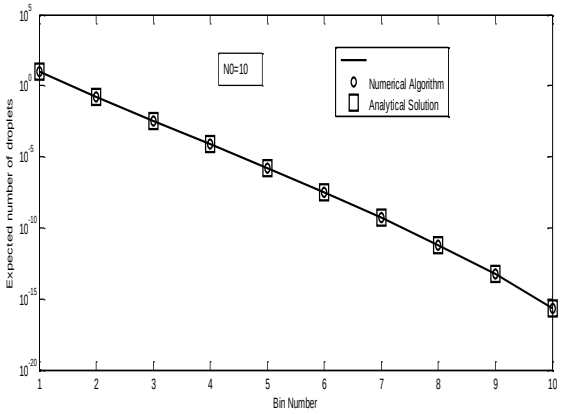

FIG. 3. For the sum kernel size distributions obtained from the analytical solution of the master equation (squares) and the numerical algorithm (circles) at *t*=500 sec. Calculations were performed with the initial condition $P(10,0,0,...,0;0)$ and the sum kernel $K(i,j)=B(x_i+x_j)$, with. B=8.82×10^2 cm^3 sec^{-1}.

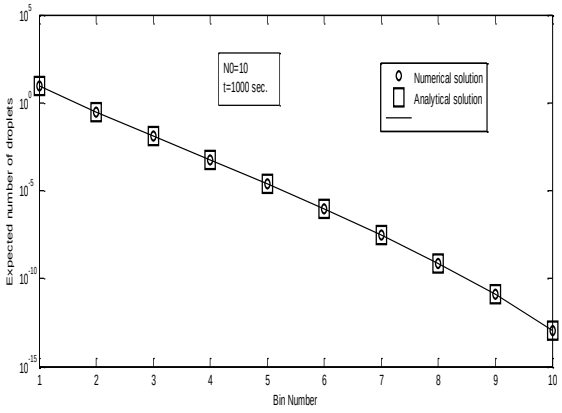

FIG. 4. Same as FIG. 3 but for *t*=1000 sec.

To conclude, some numerical results for the hydrodynamic by using the approximating polynomial proposed by Long [9]:

$$K(i,j)=9.44\times10^9(x_i^2+x_j^2) \qquad R\le 50\mu m \qquad (13)$$

Where *R* is the radius of the largest droplet, *x* is its volume in cubic centimeters and x_j the volume of the smaller droplet. The master equation was integrated with a spectrum with 10 mononers as an initial condition, and the time step was set equal to Δt=0.1 sec. The true stochastic averages at *t*=1000 sec are displayed in Fig.5, together with the mean values for each droplet mass calculated numerically from the kinetic collection equation (KCE) with kernel (13). At the large end of the spcetrum, results differ substantially from the stochastic means due to the influence of statistical fluctuations and correlations that become important in poorly mixed systems or in systems of small populations [3].

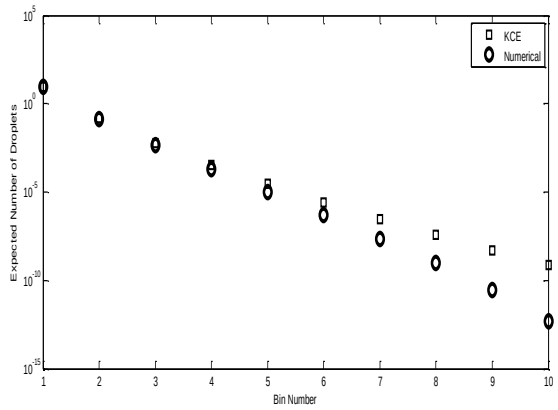

FIG. 5. For the sum kernel size distributions obtained from the numerical solution of the master equation (circles) and the KCE (squares) at *t*=1000 sec.

4. CONCLUSIONS

The full stochastic description of the growth of particles in a coalescing system is a challenging problem and an increasingly important subject for many branches of science. For finite volume systems or in systems of small populations, statistical fluctuations become important and the mathematical description rely on the master equation. Unfortunately, even for very simple models, calculating the solution of this equation is a daunting task. In an effort to solve this problem, we have introduced a new approach to numerically calculate the solution of the coalescence master equation for any type of kernels and initial conditions.

Comparison of true stochastic averages calculated numerically are in excellent agreement with analytical solutions obtained by other authors. Future work is needed in order to obtain solutions for realistic initial conditions and collection kernels modified by turbulent processes.

Acknowledgements: This study was funded by grants UACM-SECITI No. PI2011-68R and from the Consejo Nacional de Ciencia y Tecnología de México SEP-CONACYT CB-131879.

5. REFERENCES

[1]Pruppacher, H.R., and J.D. Klett, 1997: Microphysics of clouds and precipitation, Kluwer Academic Publishers.

[2]Marcus, A.H., 1968: Stochastic coalescence. *Technometrics*, Volume 10, Issue 1.

[3]Bayewitz, M.H., J. Yerushalmi, S. Katz, and R. Shinnar, 1974: The extent of correlations in a stochastic coalescence process. *J. Atmos. Sci*., 31, 1604-1614.
[4]Lushnikov, A.A., 1978: Coagulation in finite systems. *Journal of Colloid and interface science*, Volume 65, Issue 2, Pages 276–285.
[5]Lushnikov, A.A., 2004: From sol to gel exactly. *Phys. Rev. Lett*., 93, 198302.
[6]Tanaka, H., and K. Nakazawa, 1993: Stochastic coagulation equation and the validity of the statistical coagulation equation. *J. Geomag. Geoelecr*., 45, 361-381.
[7] Van Dongen, PGJ and MH Ernst, 1987: Fluctuations in coagulating systems. *Journal of statistical physics*, Volume 49, Issue 5-6, pp 879-926.
[8] Hall, M.J., 1967: Combinatorial Theory. Blaisdell Pub. Co.
[9]Long, A.B., 1974: Solutions to the droplet collection equation for polynomial kernels. J. Atmos. Sci., 31, 1040-1051.