SRC-TR-87-84

CHEMICAL PROCESS SYSTEMS LABORATORY

Applications of Condition Numbers A-Priori Analysis of Moment Method for Describing Size Distributions

by

James W. Gentry
Richard V. Calabrese

RESEARCH REPORT

CHEMICAL PROCESS SYSTEMS ENGINEERING LABORATORY

APPLICATIONS OF CONDITION NUMBERS A-PRIORI ANALYSIS OF MOMENT METHOD FOR DESCRIBING SIZE DISTRIBUTIONS

James W. Gentry
Richard V. Calabrese

A CONSTITUENT LABORATORY OF
THE SYSTEMS RESEARCH CENTER

THE UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND 20742

Particle Characterization

An International Journal
Devoted to the Measurement
and Description of Particle
and Bulk Properties in
Disperse Systems

Reprint

Applications of Condition Numbers A-priori Analysis of Moment Method for Describing Size Distributions

James W. Gentry, Richard V. Calabrese***

(Received: 26 March 1986)

Abstract

A procedure for estimating the effectiveness of algorithms which retrieve size distribution parameters from cumulative fractions or moments is described. The principle of the algorithm is to select moments, or cumulative fractions, which minimize the condition number. Extensive tests of the algorithm for a distribution consisting of the sum of two log normal distributions were carried out. This procedure can be easily extended to use different numbers and types of constituent functions. The simulations indicated that moment methods which include positive and negative moments give the best result. When the means of the constituent distribution are close, the large condition numbers indicate that no algorithm will give unambiguous values for the parameters.

1 Introduction

In a recent paper *Calabrese* et al. [1] suggest that multiple modal size distributions can be fitted with a distribution consisting of a sum of log normal distributions. The parameters are selected so that 3 N − 1 moments are matched exactly. The novelty of their approach lies in the use of condition numbers as criteria for the selection of the moments and the inclusion of negative moments. It should be noted that the peaks in the constituent distributions were widely separated with ratios between 4 and 20. In this paper we extend the analysis to two other algorithms − the use of cumulative fractions and a hybrid method combining cumulative fractions and moments. In addition we consider the case where the maxima of the constituent distributions are close and the development of criteria for initial estimates. Early applications of the algorithm suggest that it is useful for characterizing droplet distributions and atmospheric aerosols [2] but that it is not suitable for inclusion of computer codes for nuclear aerosols [3].

The basis for the analysis is that when the size distribution can be represented by a specific functional form, that is the sum of log normal distributions, the moments can be expressed unambiguously as functions of the distribution parameters. If these functions are linearized using the Newton-Rapheson method, one has a set of linear algebraic equations. An iterative procedure would then converge provided the initial estimates of the parameters were sufficiently close. We suggest that condition numbers constitute the criteria for estimating how close to the true values the initial estimates must be. If the condition number is too large ($>10^6$) then the initial estimates must be so close to the true values that the algorithm is impractical. We consider the special case where the means of two distributions are close together. Such distributions occur when unimodal distributions evolve into multimodal distributions, as for coagulation with a constant source term.

2 Condition Numbers

Condition numbers are criteria which quantitatively characterize the magnitude of the perturbation in the solution vector of a set of linear, algebraic equations due to fluctuations in the forcing function vector. Initially they were used to characterize the effect of round-off error in numerical computation. Recently they have been used in experimental design and in the a priori evaluation of inversion algorithms for experimental measurements. For these applications the forcing function represents the experimental measurement and the perturbation of the forcing function corresponds to experimental error. The coefficients of the matrix (A) are determined by the control variables of the experiment or the procedure used to analyze the experiment. These values are independent of experimental error.

The objective in the experiment design is to minimize the fluctuations in the solution vector due to experimental error. The procedure for analysis adopted here assumes that the design problem can be characterized by a simultaneous set of linear, algebraic equations:

$$A X = B . \tag{1}$$

In terms of experimental measurements B represents the experimental measurement, X the variables of interest, and A the experimental conditions or the procedures used in the inversion algorithm. ΔB and ΔX represent the perturbations in measurement and solution vectors respectively. Ideally one would prefer to have bounds for the relative error of each component in the

* Prof. *J. W. Gentry, R. V. Calabrese,* Department of Chemical and Nuclear Engineering, University of Maryland, College Park, Md. 20742 (USA).

** Present address: Laboratory for Aerosol Physics and Filter Technology at the Nuclear Center, D-7500 Karlsruhe (Federal Republic of Germany).

0176-2265/86/0310-0104 $02.50/0

solution vector. Unfortunately, one must be satisfied with bounds for the ratio of the norms of the perturbation in the solution vector to the solution vector. This results in a single number but introduces the ambiguity of the definition of the norm.
If the q-th norm of a vector is designated by $|X|_q$ one has the bound

$$\frac{|\Delta X|_q}{|X|_q} = C_q \frac{|\Delta B|_q}{|B|_q} \tag{2}$$

where the condition number C_q is defined by [4, 5]

$$C_q = |A|_q |A^{-1}|_q . \tag{3}$$

The significance of these equations is that the uncertainty in the solution to a set of algebraic equations is bounded by the product of the relative error and a parameter C_q, the condition number. Furthermore C_q depends only on the matrix A, that is the experimental conditions or the procedure for inversion, and is independent of experiment.
The principal remaining difficulty is the ambiguity due to the choice of the norm. Fortunately it is sufficient if all condition numbers show the same trend regardless of the definition of the norm. This is assured by the norm equivalence relation [6]

$$a_1 |X|_{\bar{q}} \leq |X|_q \leq a_2 |X|_{\bar{q}} \tag{4}$$

where a_1 and a_2 are constants, usually within an order of magnitude of unity, that depend only on the choice of norm. The significance of this result is that, in order to examine the effects of the condition number, it suffices to choose the norm requiring the simplest computations.
We choose the ∞ norm because it is simple to compute, requiring only the coefficients of the matrix A and its inverse. In addition, if each row of the matrix A is normalized by dividing by the sum of the absolute values in the row, the ∞ norm is then a minimum with regard to any other choice of row multiples. For a vector the ∞ norm is the component which has the largest absolute value. For a matrix it is the maximum of the sums of the absolute values of the coefficients in a row. For the normalized matrix A this maximum is one, so that the condition number is equal to the norm of the inverse matrix.
As mentioned above the condition numbers were originally developed for quantifying the effect of round-off error in the solution of algebraic equations. Recently *Farzanah* [4] showed that choosing a base set of indicator elements using condition numbers as a criterion for their choice resulted in the same base set as is currently used. This suggests an a-priori method for selection of indicator elements. The latter are a widely applied procedure for associating atmospheric aerosols with their source [7]. *Yu* tested condition numbers as a method for characterizing the performance [6] of aerosol classifiers, diffusion battery, elutriator and inertial impactor, in the determination of size distributions. He showed that the condition number proved an excellent indicator to the relative difficulty in the analysis of measurements of the size distribution. He showed that lower condition numbers were the reason why *Maigne* [8] was able to invert diffusion battery measurements using discrete fractions but not cumulative fractions of the number distribution. He found that condition numbers did not provide a test among three generic types of inversion algorithms. Recently *Kaplan* [5] developed a procedure for a priori evalution of power law correlations widely used in engineering. Her idea was to examine the data to be used in the correlation in a normalized plane where the coordinates are the logarithms of the dimensionless groups in the correlation. If the condition numbers are too large then additional measurements would be required. The method shows promise as a means of uncovering subtle dependence in ostensibly independent groups.
More directly relevant to these studies are the work of *Kaplan, Wang,* and *Gentry* [3] among others in evaluating the effect of the choice of moments on condition numbers used in the inversion of atmospheric aerosol measurements and drop break up studies. *Kaplan* [2] demonstrated that for an aerosol size distribution having the parameters suggested by *Whitby* for a trimodal distribution, the condition numbers were very sensitive to the choice of moments. The condition number varied by five orders of magnitude. *Calabrese, Bryner,* and *Gentry* [1] analyzed drop break-up data. They examined three different choices of moments. From a Monte Carlo simulation of experimental measurements they showed that the choice of moments giving the lowest condition number results in a procedure which gives parameters closer to the true distribution, that always converge, and that requires fewer data points for accurate estimates of the system parameters. The study described here is an extension of these studies.

3 Analysis

Our approach combines two ideas. The criteria for the degree of uncertainty in linear algebraic equations can be quantified by the condition number and the uncertainty in non-linear terms can be estimated from the linearization of these terms about the parameter values. This implies that one can analyze an algorithm if the steps can be expressed as a non-linear term and if these terms can be linearized. It is evident that this constitutes a necessary condition for a suitable algorithm. If the condition numbers were very large, small fluctuations in the measurements would result in large changes in the parameters which is a physically unreasonable condition.
We assume that the distribution can be characterized by the sum of log normal distributions. Specifically, in the examples below, the test distribution consists of two log normal distributions.

$$f(x) = \left[\frac{c_1}{\sqrt{2\pi}\,\sigma_1} \operatorname{Exp} - \frac{\operatorname{Ln}^2(x/x_1)}{2\sigma_1^2} + \frac{c_2}{\sqrt{2\pi}\,\sigma_2} \operatorname{Exp} - \frac{\operatorname{Ln}^2(x/x_2)}{2\sigma_2^2} \right] \frac{\mathrm{d}\operatorname{Ln} x}{\mathrm{d}x} \tag{5}$$

where c_i, X_i, and σ_i are the number fraction, mean size and standard deviation of the respective distributions. If the distribution function is normalized, one has:

$$c_1 + c_2 = 1 . \tag{6}$$

We refer to this equation as the normalization condition. It is used as one of the equations in each of the methods. To specify the distribution, unambiguously, it is necessary to specify five additional equations. The specification of these conditions defines the inversion algorithm.
A third possibility is a combination of moments and cumulative fractions for up to five different values of p. If the distributions are log-normal, then the moments or cumulative fractions are given in closed form. Next the equations are expanded in a

Taylor series and truncated after the first term. The equations for the p-th moment become

$$M(p) = c_1 X_1^p \text{ Exp } \frac{p^2\sigma_1^2}{2}\left[\frac{\Delta c_1}{c_1} + p\frac{\Delta X_1}{X_1} + p^2\sigma_1\Delta\sigma_1\right] + c_2 X_2^p \text{ Exp } \frac{p^2\sigma_2^2}{2}\left[\frac{\Delta c_2}{c_2} + p\frac{\Delta X_2}{X_2} + p^2\sigma_2\Delta\sigma_2\right] + c_1 X_1^p \text{ Exp } \frac{p^2\sigma_1^2}{2} + c_2 X_2^p \text{ Exp } \frac{p^2\sigma_2^2}{2} \quad (7)$$

or

$$M(p) = G_1\left(y_1 + py_2 + \frac{p^2}{2}y_3\right) + G_2\left(y_4 + py_5 + \frac{p^2}{2}y_6\right) + G_1 + G_2 . \quad (8)$$

The algebraic equations have the form

$$A_{ji}y_i = B_j \quad (9)$$

where B_j is defined by

$$B_j = m(p_j) - G_1(p_j) - G_2(p_j) . \quad (10)$$

A similar relation can be defined for the cumulative fractions. The condition numbers are computed from the coefficients A_{ji}. We choose the ∞ norm and we normalize the equation so that the sum of the absolute values are one. The norm of a matrix is the maximum value of the sum of the absolute values of the coefficients in each row.

To illustrate the method we consider a distribution consisting of only a single log-normal distribution. The area ($p = 2$) and volume ($p = 3$) moments are specified. The resulting matrix of coefficients is

$$\begin{bmatrix} 1 & 0 & 0 \\ G(2) & 2G(2) & 4G(2) \\ G(3) & 3G(3) & 9G(3) \end{bmatrix} . \quad (11)$$

After normalization the coefficients are

$$\begin{matrix} 1 & 0 & 0 \\ .143 & .286 & .571 \\ .077 & .231 & .692 \end{matrix} \quad (12)$$

which has the inverse

$$\begin{matrix} 1 & 0 & 0 \\ -.83 & 10.5 & -8.67 \\ .17 & -3.5 & 4.33 . \end{matrix} \quad (13)$$

The condition number is defined as

$$C_\infty = |A|_\infty \cdot |A^{-1}|_\infty = 1 \cdot (.83 + 10.5 + 8.67) = 20 . \quad (14)$$

The choice of other values of p may result in a lower condition number but at the cost of greater experimental error.

Two additional points should be made. If the algorithm uses only moments and if the standard deviations are the same for the two component distributions, then the condition number is independent of the standard deviation. Secondly, it is sufficient to specify the ratio of (X_2/X_1) rather than both values. When the standard deviations are equal the condition numbers for any algorithm has two parameters, the fraction of the particles in the larger size distribution c_2 and the ratio of the mean values of the two distributions.

4 Description of Algorithms

Three classes of algorithms were used in the study. They are designated in Table 1. They are algorithms based only on the moments, designated with an M, algorithms based only on cumulative fractions, designated with a C, and algorithms containing both moments and cumulative fractions, designated with an H. The condition numbers do not weigh experimental errors and it is possible that an algorithm which is non-feasible, because the moments or cumulative fractions cannot be determined with any degree of confidence, yields the lowest condition numbers. In practice the choice of algorithm must balance small condition numbers with small experimental uncertainties.

Table 1: Description of inversion algorithms.

Method	Moment	Cumulative fractions			Weight
1. $M(-3; 3)$	0, −1.5, −3, 1, 2, 3	–			
2. $M(-1.5; 1.5)$	0, −0.75, −1.5, 0.5, 1.0, 1.5	–			
3. M (0, 3)	0, 1, 1.5, 2, 2.5, 3				
4. $C(0)$	0	0.17	0.33	0.50	0
		0.67	0.83		0
5. $C(0; 3)$	0	0.33	0.67		0
		0.25	0.50	0.75	3
6. $H(0; 3)$	0,3	0.33	0.67		0
		0.33	0.67		3

The particle size distributions consist of two log-normals in which the ratio of sizes and the fraction of particles in the distribution with the larger mean (c_2) are parameters. The standard deviation was 0.8 for both distributions. This is a moderately broad distribution, typical of atmospheric and nuclear aerosols. The values of c_2 ranged from 0.5 to 0.001 and the size ratio ranged from 2 to 100. These are the regions of practical applicability, although we have restricted our discussion to sizes from 5 to 20 in three size rations, 6, 10, and 18.

For the algorithms, designated by M, based on moments, it is necessary to specify five different moments. The equations were linearized and the condition numbers were computed. In practice all the moments except area ($p = 2$) and volume ($p = 3$) are hypothetical. For an analysis based on photographs or electron micrographs of the particles, hypothetical moments, including negative moments, are as meaningful as volume or area. When using hypothetical moments there is less error in computing moments than in computing cumulative fractions.

Algorithms based on cumulative fractions are designated with C. We considered only those algorithms which consist of cumulative fractions with one or more experimentally

measurable weight, that is $p = 0$, 2, or 3. We used equally spaced fractions. The condition number will vary with the choice of fractions but numerical simulations to quantify this effect are unlikely to be useful.

The third type of algorithm are hybrid methods, designated by H. These include both moments and cumulative fractions. They provide a fixed reference, a distinct advantage compared to the cumulative fraction algorithms. They work best when the constituent distributions have small standard deviations.

5 Numerical Results

Three different classes of algorithms were investigated. These were the method of moment, the method of cumulative fractions and a hybrid method which includes both moments and cumulative fractions. The distribution consisted of the sum of two log normal distributions containing six parameters. In all cases the normalization relation that the sum of the fractions in the two distributions is equal to one is one of the six equations. The other five consist of moments, cumulative fractions, or combination of the two. In the numerical examples below the values of the standard deviations are equal. This restraint is not necessary for the analysis. We assume that the ratio of the mean sizes for the two distributions are between 5 and 20 and the fraction of particles in the distribution with the larger mean ranges from 0.5 to 10^{-3}. For all methods the condition numbers were large with values near to 10^4. In contrast the condition number for a single log-normal distribution is 10. This indicates that the measurements must be significantly better to obtain the parameter for two distributions unambiguously.

The methods which gave the most satisfactory results are the moment method, with moments ranging from -3 to 3 (Figure 1) and the hybrid method with two moments 0 and 3 (Figure 2). The methods based on cumulative fractions alone (Figure 3) have such high conditions numbers that their utility is doubtful. In these curves the condition number is plotted as a function of c_2 for three size ratios (6, 10, 18). The condition numbers for methods based on moments increase with decreasing number fraction, that is the fraction of the particles in the distribution with the larger mean. This effect can be mitigated by choosing larger moments. However, it suggests that this class of algorithms would be more effective when the number fraction is greater than 0.02. The larger the ratio of mean sizes the more effective is the algorithm. Satisfactory inversions should be obtained so long as the number fraction is greater than 0.01. The

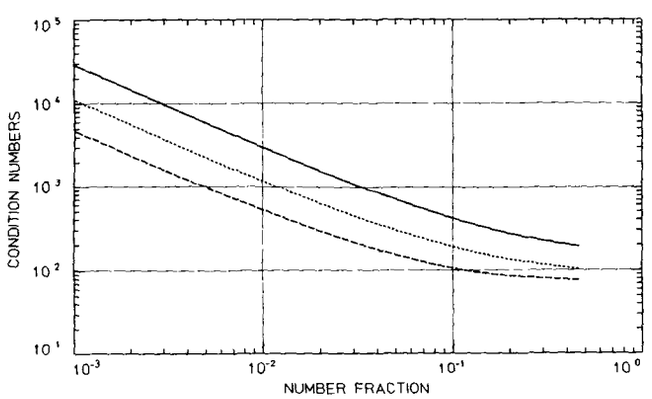

Fig. 1: Condition number: moments (-3; 3). sigma = 0.4, size ratio = -6, ··· 10, --- 18

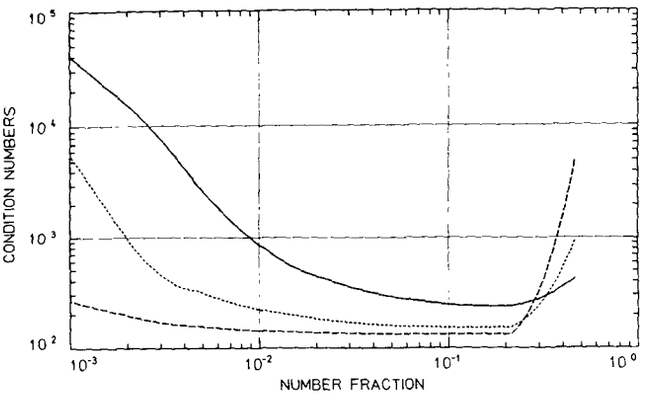

Fig. 2: Condition number: hybrid method (0; 3). sigma = 0.4, size ratio = -6, ··· 10, --- 18

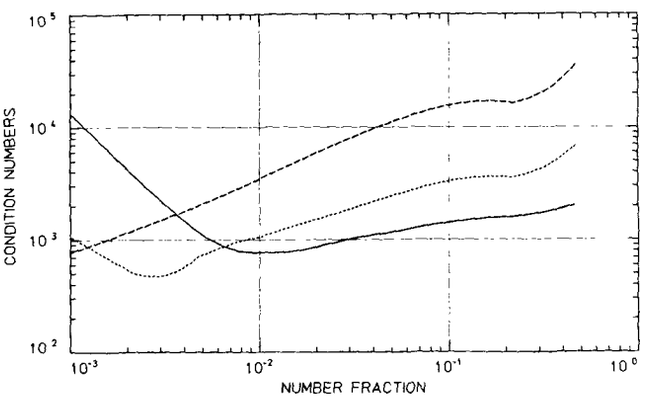

Fig. 3: Condition number: cumulative fractions (0; 3). sigma = 0.4, size ratio = -6, ··· 10, --- 18

condition numbers are independent of the standard deviations when the standard deviations are equal.

The hybrid method consisting of two cumulative fractions for both the 0 and 3 moment as well as the moment for $p = 3$ gives a small condition number. When the means are close together, or when the distributions are broad, this algorithm is less effective. The flat section for a wide range of number fractions is characteristic of this algorithm.

It is interesting to contrast these results with the condition numbers for cumulative fractions at two different weights ($p = 0$ and $p = 3$). The difference in the algorithms is that the cumulative fractions have no absolute measurements while the hybrid method includes the moment ($p = 3$). There is a difference of one order of magnitude for a common sigma (0.8). This demonstrates the advisability of incorporating at least some absolute measurements in the inversion algorithm. Again there is little dependence of condition number with number fraction. Nevertheless the principal result is that cumulative fractions do not constitute the basis for a satisfactory inversion algorithm.

For comparison of algorithms in which only the moments are computed, two rules can be formulated:

1. The greater the separation of the moments, the smaller the condition number.
2. The larger the maximum of the absolute values of the moments, the greater the experimental error. When the particles are counted and sized individually, all moments ($p \neq 0$) are hypothetical. One would expect the experimental error in $M(-3; 3)$ to be the same as $M(0; 3)$ or $H(0; 3)$, and this error to be greater than for $M(-1.5; 1.5)$. The condition number, which is unaffected by experimental error, should be smaller for $M(-3; 3)$ than for $M(-1.5; 1.5)$.

The ratios of condition numbers $M(0; 3)$ to $M(-3; 3)$ and $M(-1.5; 1.5)$ to $M(-3; 3)$ are plotted in Figures 4 and 5. The significant results are:

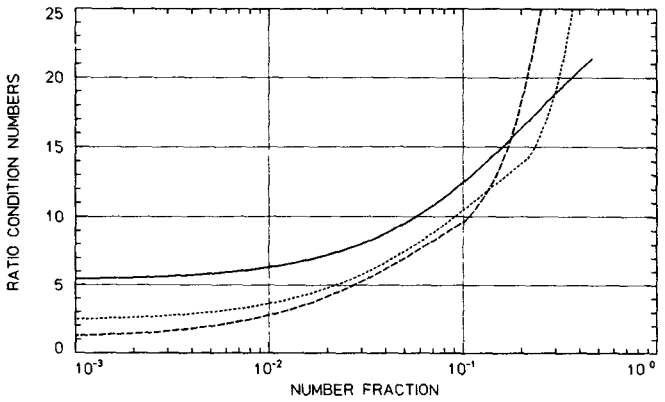

Fig. 4: Comparison moments (0; 3) with moments (3; 3). sigma = 0.4, size ratio = −6, ··· 10, --- 18

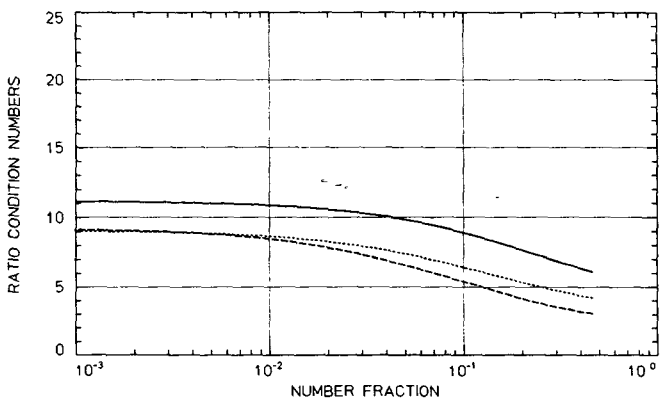

Fig. 5: Comparison moments (−1.5; 1.5) with moments (3; 3). sigma = 0.4, size ratio = −6, ··· 10, --- 18

1. There is a penalty of less than one order of magnitude in using $M(-1.5; 1.5)$ rather than $M(-3; 3)$.
2. This is less than the penalty for using $M(0; 3)$ when c_2 is greater than 0.01 which is the range of greatest interest.
3. As the size ratio increases, the efficiency of all three methods $M(-3; 3)$, $M(-1.5; 1.5)$ and $M(0; 3)$ converge.

Choosing an inversion algorithm with negative, albeit hypothetical, moments results in more efficient algorithms. It is probable that the reduced effect of experimental error in $M(-1.5; 1.5)$ compared to $M(-3; 3)$ outweights the slightly larger condition numbers.

We next consider the choice of algorithm. In Figures 6 and 7, the algorithms based on cumulative fractions $C(0)$ and hybrid methods $H(0; 3)$ are compared to the moment method $M(-3; 3)$. The ratio of condition numbers for two methods is plotted as a function of c_2 for three different size ratios. The standard deviation of the distribution was 0.4. This comparatively narrow distribution results in smaller condition numbers for $H(0; 3)$ and $C(0)$ but does not affect the results of the moment methods.

The hybrid method is more effective than the moment method for large differences in the mean size. Effectively the distribution consists of two, disjoint log-normal which are analyzed separately by the algorithm. The cumulative fraction method results in condition numers that are 1 to 3 orders of magnitude higher than the moment method.

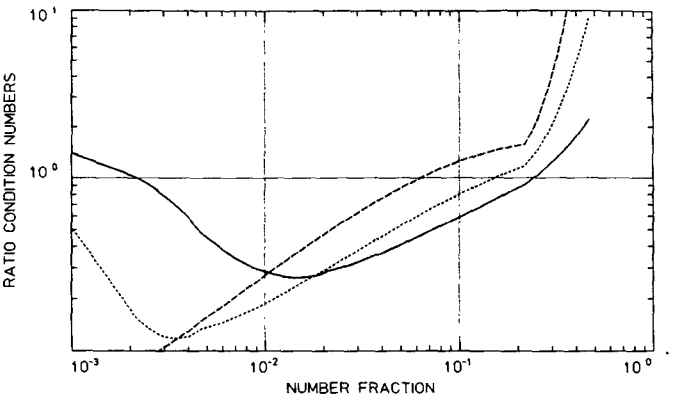

Fig. 6: Comparison hybrid method (0; 3) with moments (3; 3). sigma = 0.4, size ratio = −6, ··· 10, --- 18

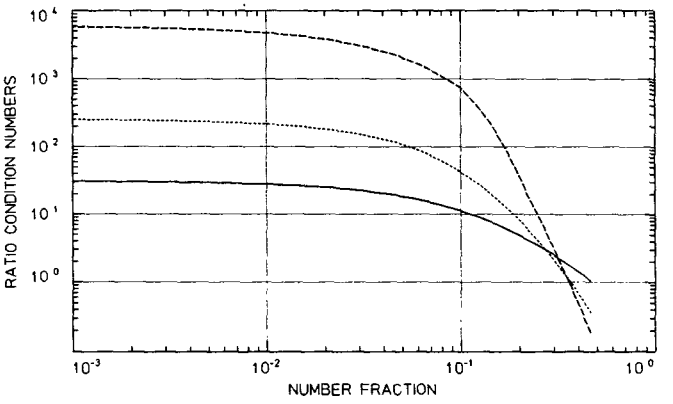

Fig. 7: Comparison cumulative fractions (0) with moments (3; 3). sigma = 0.4, size ratio = −6, ··· 10, --- 18

In summary the moment method is the better of the three algorithms. The hybrid method will give better results when the size ratio of the means is large. The cumulative fraction method can be justified only when the fractions of the two log-normal distributions is near to 0.5.

6 Small Size Ratios

As the size ratio approaches one, the distribution collapses into a single log-normal distribution. The set of six equations are no longer independent and the condition number becomes unbounded. Thus all of the algorithms fail. The important point is whether the algorithms are useful for small size ratios which are different from one and at what size ratio the algorithms fail. This information is significant in the use of specific functional forms in solutions of evolution equations. For these equations a number of related solution procedures have been suggested. In one procedure a set of ordinary differential equations, the moments of the population balance equation, are solved by Runge-Kutta methods. The parameters for the distribution, a log normal function, are obtained from the moments. In an alternative approach the distribution function is expanded in terms of the time derivatives of its parameters. These time derivatives are chosen to minimize a functional, that is the absolute value of the residual of the population balance, integrated over all particle sizes. Each of these approaches could be extended to multi-modal distributions provided the set of 3 N equations can be solved when the means are close together.

Typical values for the condition number are plotted as a function of size ratio for three algorithms ($M(-3;3)$, $H(0;3)$, and $C(0)$) in Figures 8, 9, and 10. The size ratio varies from nearly, 1.0 to 2.5. The three curves correspond to different values of c_2 (0.464 -. 0.215 - - -, and 0.100 - -). Each of the methods give better results, that is lower condition numbers, as c_2 approaches 0.5. We chose low values of standard deviations ($\sigma_i = 0.4$) so that the hybrid method and cumulative fractions could be tested under favorable conditions. The principal result is that none of the methods are satisfactory for a size ratio less than 1.5. The condition numbers are $>10^5$. The ratio where the condition numbers ($c_2 = 0.464$) are 10^7 are 1.17, 1.23, and 1.18 respectively. The moment method is marginally the best of the three while the hybrid method is marginally the worst. However, each of the methods is unsatisfactory.

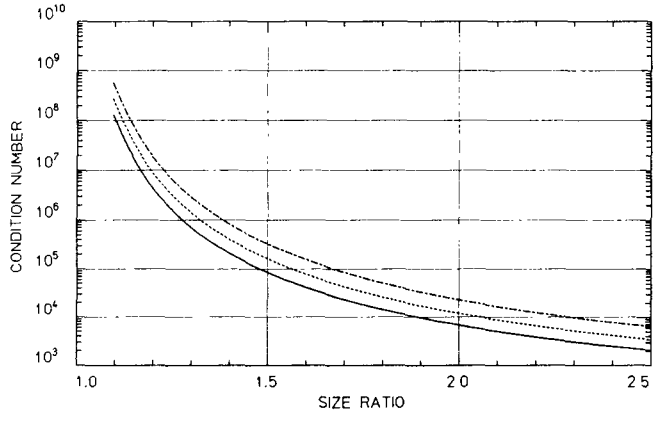

Fig. 8: Condition number: moments (−3; 3). sigma = 0.4, fractions 0.46, 0.22, 0.10

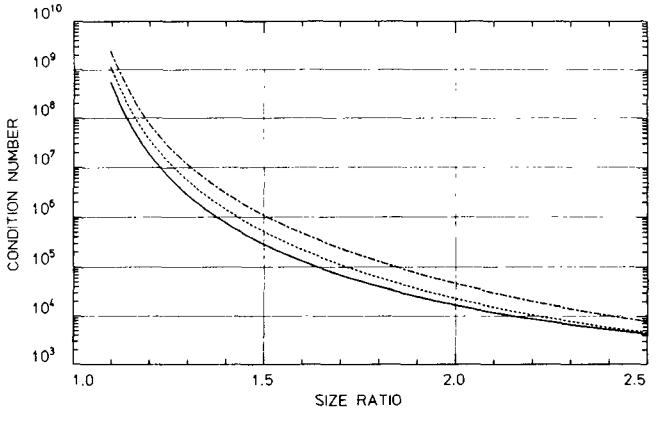

Fig. 9: Condition number: hybrid method (0; 3). sigma = 0.4, fractions 0.46, 0.22, 0.10

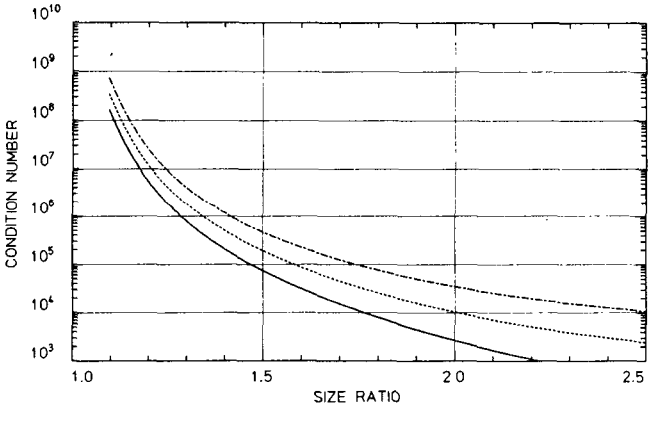

Fig. 10: Condition number: cumulative fraction (0). sigma = 0.4, fractions 0.46, 0.22, 0.10

7 Initial Estimates

Each of the algorithms require initial values for the distribution. Since the equations are characterized by large condition numbers, it is critical that the initial trial values be as close to the final values as possible. Our procedure is based on two ideas – we examine the behavior of the moments under the condition that the distribution with the smaller size affects only the negative moments and the particles in the distribution with a larger mean size influences only the positive moments. Secondly we restrict the parameters to feasible values. With our method we establish which of the c_i is greater than 0.5. For this distribution we establish an upper bound for σ_i and x_i. We obtained no restriction on the second distribution.

The first and critical step in the algorithm is to choose the constituent distribution where $c_i > 0.5$. Estimates of the standard deviation and mean size are obtained for two values of p. (Usually $p = 2$ and $p = 3$). The equations are

$$\sigma_2(p) = \ln \frac{M(p+1)\,M(p-1)}{M^2(p)}$$

$$x_2(p) = \frac{M(p+1)}{M(p)} \operatorname{Exp} - \frac{(2p+1)}{2}\sigma_2^2$$

$$\sigma_1(p) = \ln \frac{M(-p+1)\,M(-p-1)}{M^2(-p)}$$

$$x_1(p) = \frac{M(-p-1)}{M(-p)} \operatorname{Exp} - \frac{(2p+1)}{2}\sigma_1^2 . \qquad (15)$$

These equations are exact when the two distributions are disjoint.

Next we consider the auxiliary functions defined by:

$$\Phi_i(1) = \left[\frac{\sigma_i(p+1)}{\sigma_i(p)}\right] - 1$$

$$\Phi_i(2) = \left[\frac{x_i(p+1)}{x_i(p)}\right] - 1 . \qquad (16)$$

The component where $c_i > 0.5$ has the property that $\Phi_i(1)\,\Phi_i(2) > 0$. This means that the Φ_i have the same sign. There are two potential complications: both components satisfy the criteria or neither satisfy the criteria. In all of our test cases these complications did not arise, although we did not examine the case where $c_i \equiv 0.5$.

Once the value of c_i is found there are two inequalities:

$$\sigma_i \le \text{Minimum } \sigma_i(p),\ \sigma_i(p+1)$$

$$x_i \le \text{Minimum } x_i(p),\ x_i(p+1);\ i = 1$$

$$x_2 \ge \text{Maximum } x_2(p),\ x_2(p+1);\ i = 2 . \qquad (17)$$

The algorithm was tested for a large number of test cases. Convergence was obtained for all cases where $c_i < 0.95$ and the size ratio was 2.2 or greater. The method frequently failed when the size ratio was 1.5 and c_i was greater than 0.9. It returned a log-normal distribution. By failure, one means that repeated iterations did not converge to the true values or to two distributions which could not be distinguished from the true distribution.

Cases using distributions with different σ_i give a more demanding test than when the σ_i are equal.

8 Conclusions

Condition numbers were used as an a-priori criterion for evaluating algorithms used to fit multi-modal size distributions. Three classes of algorithms, those based on moments, on cumulative fractions, and a hybrid method which include both cumulative fractions and moments were examined. The calculations examined the effect of the relative number concentrations between the two constituent distributions and the effect of the ratio of the mean sizes of the two distributions.

It was concluded that the preferable algorithms were those based on moments. Except for number fractions near to 0.5, the cumulative fractions gave substantially higher condition numbers. The hybrid methods gave the best result when the two distributions were widely separated and when c_1 was greater than 0.98.

For algorithms based on moments, those containing negative moments were superior to those containing only positive moments (provided the spread between the maximum and minimum value were the same). The cost in terms of higher condition numbers, for algorithms where $p_{max} = 1.5$ compared to those with $p_{max} = 3.0$, was less than one order of magnitude. Because of decreased experimental error this algorithm is recommended.

None of the algorithms give satisfactory results when the size ratio is less than 1.5. We believe that this precludes the use of multimodal distributions in characterizing evolving particle distributions.

9 Appendix (Step by Step Procedure in Evaluating Distributions)

Step 1: Select the functional form of the trial distribution function and determine the number of parameters. In our study we used the sum of two log-normal distributions which contain six parameters.

Step 2: Write the integrated expressions for the moments and weighted cumulative fraction in as compact form as possible. For log-normal distributions these variable can be expressed in closed form.

Step 3: Select the inversion algorithm to be tested. A total of $N - 1$ moments or weighted cumulative fractions are chosen.

Step 4: Expand the equations in Taylor series about the parameters. Retain only the linear terms. This results in a set of N simultaneous equations.

Step 5: Normalize the resulting matrix so that the sum of the elements in each row is unity. Compute the inverse of the matrix. From the inverse determine the ∞norm. This is the condition number.

Step 6: Determine the condition number for a multiple dimensioned lattice. Since the dimensionality of the lattice is $N - 1$ it is necessary to make a judicous selection. We were able to divide through by the minimum size and we assumed that the σ_i were comparatively less important. As a result we retained two parameters, the size ratio and the relative fraction.

Step 7: The algorithm which has the smallest condition number in the range of interest is preferable. It is less sensitive to errors in initial estimates. This analysis neglects experimental error.

10 Symbols and Abbreviations

Upper Case Letters

A, A_{ij}	matrix, matrix elements
A^{-1}	inverse of matrix A
B, B_j	forcing function, components of vector
$C(\,;\,)$	designation for inversion algorithm, based only on cumulative fractions
C_q^*	condition number defined for the q-th Norm
$G_i(P)$	auxillary function equal to the i-th contribution for the p-th moment
$H(\,;\,)$	designation for inversion algorithms, based on both cumulative fractions and moments
$M(\,;\,)$	designation for inversion algorithm, based only on moments
N	number of parameters in distribution function
X, X_i	solution vector, components of vector perturbation operator, X is the perturbation of the variable X

Lower Case Letters

c_i	relative frequencies or fractions of the i-th lognormal components
f	frequency or distribution function
$m(p)$	the p-th moment
p	index indicating order of moment
q	index indicating specific norm
x_i	logarithmic mean size of the i-th lognormal component
$x_i(p)$	estimate or trial value for x_i (defined above) based on p-th moment
x	size parameter
a_i	numerical constants
Φ_i	auxillary functions used in iterative steps to obtain distribution parameters
σ_i	standard deviation of the i-th lognormal component
$\sigma_i(p)$	estimate or trial value of the i-th lognormal component

Miscellaneous Symbols

A_q, B_q the i-th norm of the matrix A or the vector B.

11 References

[1] *R. V. Calabrese, N. P. Bryner, J. W. Gentry:* Analysis of Drop Size Distributions Produced in Dilute, Agitated Liquid-Liquid Systems. Proc. 1. World Congress Particle Technology. Nürnberg (1986), Part IV, pp. 585-601.

[2] *C. R. Kaplan, R. V. Calabrese, J. W. Gentry:* Condition Numbers: Application to Correlations, Distribution Analysis and Experiment Location. AAAR Annual Meeting, Albequerque, Nov. 1985.

[3] *C. R. Kaplan, J. Wang, J. W. Gentry:* Condition 'Numbers, Monte Carlo Simulations, and Filter Clogging'. GAeF Annual Meeting, Garmisch-Partenkirchen, September 25-27, 1985.

[4] *F. Farzanah, C. R. Kaplan, P. Y. Yu, J. Hong, J. W. Gentry:* Condition Numbers as Criteria for Evaluation of Atmospheric Aerosol Measurement Techniques. Environ. Sci. Technol. *19* (1985) 121.

[5] *C. R. Kaplan:* M. Sc. Thesis, University of Maryland, College Park, 1985.

[6] *P. Y. Yu:* Ph. D. Thesis, University of Maryland, College Park, 1983.

[7] *G. E. Gordon:* Receptor Models. Environ. Sci. Technol. *14* (1980) 792-800.

[8] *J. P. Maigne, P. Y. Turpin, G. Madelaine, J. Bricard:* Nouvelle Methode de Determination de la Granulometrie d'un Aerosol au Moyen d'un Batterie de Diffusion. J. Aerosol Sci. *5* (1974) 339-355.