LBL-8875 c.2
UC-32

A NUMERICAL INVESTIGATION INTO THE PROPERTIES OF FUZZY NUMBERS

RECEIVED
LAWRENCE
BERKELEY LABORATORY

APR 11 1979

LIBRARY AND
DOCUMENTS SECTION

Barbara A. Cerny

December 1978

Prepared for the U. S. Department of Energy
under Contract W-7405-ENG-48

FOUR-WEEK LOAN COPY

This i
which may be borrowed for two weeks.
For a personal retention copy, call
[redacted]

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY

LBL-8875 c.2

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

LBL 8875

A NUMERICAL INVESTIGATION INTO THE PROPERTIES OF FUZZY NUMBERS

Barbara A. Cerny

Information Methodology Research Project
Lawrence Berkeley Laboratory
University of California
Berkeley, CA. 94720

December 11, 1978

Abstract

The concepts of fuzzy numbers and fuzzy arithmetic are introduced and some of their numerical properties are investigated. Fuzzy numbers are presented as extensions of interval numbers, or possibility distributions between upper and lower bounds. These distributions are described in the literature as being shape invariant under addition and subtraction and "nearly shape invariant" under multiplication and division. This shape invariance is assessed through Fortran IV programs that multiply and divide fuzzy numbers by two different methods: 1) a point by point manipulation analogous to the convolution of probability densities that gives an exact result, and 2) an approximation to this result obtained by operations on a parameterized form of the possibility distributions. Since this approximation gives tolerable results given the basic intrinsic vagueness of the numbers, sequences of operations on these parameterized fuzzy numbers are examined via a Fortran IV program. The resultants are discussed in terms of error propagation.

Introduction

When one attempts an analysis of a complex system, invariably there are some variables which are imprecisely known. Traditionally, this inexactness is treated as a form of randomness and statistics are used to try and understand the behavior of the system. However, there are situations where the inexactness is not due to randomness but rather due to imprecisely defined variables, such as occur in questionnaire data or any form of data which arises from human judgement, reasoning or perception. In order to handle this imprecision, L.A. Zadeh (1965,1973) introduced the concepts of approximate reasoning and fuzzy sets. In his approach, human reasoning does not approximate an exact logical or mathematical process; rather people use inexact concepts directly to describe and make sense of the world around them. A mathematical representation for this reasoning is provided by fuzzy set theory; it is a way to describe and analyse this form of imprecise data. While the primary aim of this paper is to examine the potential use of fuzzy arithmetic in data analysis, secondarily it should serve as a simple introduction to the concept of fuzzy sets.

Fuzzy Set Theory

People communicate through phrases such as "go a short distance then make a hard right". We all have a feeling for what constitutes a "tall man", "sort of big" or "about 3", but the boundaries of these classes of objects are vague. These ideas date back to the Greeks who

expressed them in a series of paradoxes such as "How many hairs must be plucked from a man's head before he is considered bald?" or "Where does bald end and not bald begin?" The boundary of the term bald, the transition between applicability and non-applicability of it is vague, as it is in much of language.

"Bald" or "tall" or "about 3" are called linguistic variables which are labels for fuzzy sets and in fuzzy logic these coexist and interact with traditional numerical variables. The underlying idea is that membership in a set such as "old" is not a step function with values of 0 corresponding to "not belonging" and 1 corresponding to "belonging", but rather a grade of membership in the interval (0 ... 1). As an example, consider the proposition "John is old". If OLD is defined to be the set { 20... 100} and John is 30, then he could be OLD to the degree .2 while at 90 he could be OLD to the degree 1.0. That is, each age of John's has a compatibility with the set OLD or a degree of membership in it. This leads to a distribution for fuzzy variables that Zadeh (1977) describes as a possibility distribution rather than a probability distribution.

These ideas have great intuitive appeal for data analysis purposes since so much of so many systems in the real world are fuzzy. The difficulty with using them is twofold; first the degrees of membership or the distributions must be assigned and then they must be analysed. The assignment of grades of membership, or possibilities, differ from determining probabilities in fundamental ways. First, possibility can be subjectively assigned by people in a given context; the

uncertainties in assignment are due to an intrinsic vagueness and arbitrariness rather than, as in a probabilistic event, a lack of knowledge about the future. Although it is possible to arrive at a degree of membership by questioning many people, the assumption of repeated sampling is not necessary as it is in statistics.

Assuming we could assign grades of membership to variables, their analysis is the next problem in a real world context. From the theoretical point of view, a body of literature is being developed that allows the mathematical manipulation of fuzzy sets. From the computational point of view less work has been done, so it is difficult to get a "feeling" for the behavior of these variables. This paper addresses itself to this problem by investigating the computational behavior of fuzzy numbers. For simplicity, we will not deal with linguistic variables such as OLD, but only with numerical variables of the "about 3" sort. The techniques explored, however, are eventually applicable to both since their form is essentially the same.

Since the assignment of grades of membership is always subjective and somewhat arbitrary, it seems logical to first investigate the behavior of arithmetic operations with these numbers to see if they are well-behaved and interpretable. Only then does the issue of assigning grades of membership become germaine. Since this paper is an introduction to the concepts of fuzzy arithmetic, the thrust is simplicity and comprehensibility so that a potential user can get a "feeling" for the behavior of these numbers. There is no pretense of

rigor; the cited references can be referred to for that purpose.

Interval Numbers and Fuzzy Numbers

Fuzzy Arithmetic is an extension of interval arithmetic (Moore, 1966) which arises when a number is considered to have an exact value and an error bound, or an upper and lower bound to the exact result. An interval number is an ordered pair of real numbers $[a,b]$ with $a \leq b$ or a set of real numbers $\{x \mid a \leq x \leq b\}$.

In interval analysis, all values between a and b are equally possible or probable. A fuzzy number is an extension of this concept in that a number X = "about 2" takes a functional form between a and b (see figure 1). This function is described by the degree of compatibility μ of each number in this range with "about 2". These compatibilities are represented by values in the set $\{ 0. \ldots 1.\}$ with 0 representing no compatibility with "about 2" and 1 representing the maximum compatibility or grade of membership in "about 2". In a simple case, let "about 2" range from 1 to 3 as in figure 1. The degree of compatibility at X = 1 is 0, so $\mu [1] = 0$. Likewise,

$$\mu [1.5] = .5$$

$$\mu [2.0] = 1.$$

$$\mu [2.5] = .5$$

$$\mu [3.0] = 0.$$

This is often expressed in the form $\mu (x)/x$, or the value of the

degree of compatibility at some X. In the above example, the degree of compatibility at X = 1.5 can be written as $\mu(1.5)/1.5 = .5/1.5$. To generalize this, a fuzzy number A can be considered to be the union of all the values for which $\mu_A > 0$. So

$$A = \bigcup \mu_A(X)/X$$

or

$$A = \Sigma \mu_A(X)/X.$$

where Σ stands for union. If A is a continuous function, the integral sign is used to denote union as in

$$A = \int_{X \in R} \mu_A(X)/X.$$

In figure 1, A = "about 2" is shown as a continuous function that can be represented by a triangular distribution as

$$\underset{\sim}{2} = \int_1^2 (X-1)/X + \int_2^3 (3-X)/X$$

(Mizumoto & Tanaka (1976)). Here, + stands for union not arithmetic sum. The range 1 to 3 is called the support of "about 2". In general, if Γ_A is defined to be the support, then

$$\Gamma_A = \{X | \mu_A(X) > 0\}.$$

The shape of a function can be determined in various ways. For example, B in Figure 1 could have arisen by asking a person a series of questions such as

"What is the most possible value of the number?" [MU = 1.]

"What is the point below/above which it is not possible for the number to lie?" [MU = 0.]

"At what point below/above the mean is it equally possible for the number to be greater than or less than?" [MU = .5]

Closer to Zadeh's possibility postulate (1977) but more difficult to carry out in practice are questions about the degree to which a number fits ones perception of "about 2". If you say X is "about 2", then the possibility that X has a specific value, say 2.5, is equal to the grade of membership of 2.5 in "about 2". More formally, the proposition "X is 'about 2' " induces a possibility distribution which associates with each value u between 1 and 3 the possibility that u could be a value of X. That is, the possibility of X taking a value u is equated to the grade of membership of u in the fuzzy set "about 2" or the degree of compatibility of X with "about 2". We could say based on Figure 1, A:

Possibility[X=u] = Poss[X=1.0] = 0.

Possibility[X=1.5] = .5

Possibility[X=2.0] = 1.0 etc.

If one can justify, or arbitrarily assume a functional form, a number can be fuzzified with only the answers to these three questions. Other examples of building functions in this form occur in the psychological literature where subjects are asked to give the degree of compatibility between a form and a standard, as the grade of membership of bat in the category bird.

In general, it is possible to fuzzify any domain of mathematical reasoning that is based on set theory. One replaces the precise concept that a variable has a value with the fuzzy concept that a variable has a degree of membership.

Another way of representing fuzzy numbers is in terms of bandwidth at some value of μ, as in C in Figure 1. Dubois and Prade (1978) use this form and express a number in terms of its left bandwidth, peak and right bandwidth. In this paper, this will be represented by $C = (\alpha, C, \beta)$. It is most convenient to use the bandwidth at $\mu = .5$ due to the subjective nature of the methods used to determine points on distributions. This is true even in the case of fuzzy numbers that are normally distributed. For instance, let F be a continuous function normally distributed centered on 0 with a SD of 1. In the fuzzy system, $\mu(1)$ or $\mu(-1) = .6067^{1.}$, a value difficult to obtain. It is interesting to compare this to P(X), the probability of X. For a continuous probability density, if $X = -1$, $P(X) = 0$. If F were a discrete probability density, then, of course, P(X) could be non-zero since P(X) is the number of occurrences of X compared to the total sample space. This simple example shows a fundamental difference between numbers conceived in a fuzzy versus a statistical framework. A fuzzy number can be continuously distributed yet possess a value at a discrete point, while in the analogous probability case, P(X) is defined as the area under the normal curve between two limits.

An interesting extension of these ideas comes from Nahmias' (1978) axiomatic approach in direct analogy to the sample space model of probability theory. He shows that for normal fuzzy variables, the bandwidths add, in contrast to uncorrelated random variables where the variances add. This could suggest a methodology for testing whether a probabilistic model or a fuzzy model is more appropriate in a given context.

Interval Arithmetic and Fuzzy Arithmetic

In general if $*$ is one of the symbols $+,-,\cdot,/$, arithmetic operations on intervals are defined as

$$[a,b]*[c,d] = \{x*y \mid a \leq x \leq b, c \leq y \leq d\}$$

except that $[a,b]/[c,d]$ is not defined if $0 \in [c,d]$. This definition is set-theoretic, hence the result of any operation is the set of sums, differences, products and quotients of pairs of real numbers one from each interval. In terms of formulae for endpoints,

$$[a,b]+[c,d] = [a+b,c+d]$$

$$[a,b]-[c,d] = [a-d,b-c]$$

$$[a,b]\cdot[c,d] = [\min[ac,ad,bc,bd],\max[ac,ad,bc,bd]]$$

and if $0 \notin [c,d]$, then

$$[a,b]/[c,d] = [a,b]\cdot[1/d,1/c]$$

The degenerate interval $[a,a]$ is the real number a. These numbers do not necessarily distribute, for example

$$[1,2][[1,2]-[1,2]] = [1,2][-1,1] = [-2,2]$$

$$[1,2][1,2]-[1,2][1,2] = [1,4]-[1,4] = [-3,3].$$

This is rare in algebraic systems, and hence Moore (1966) notes interval arithmetic is a little studied system.

Fuzzy arithmetic is approached from two points of view. The first is a point by point computation of values within in each distribution corresponding to a convolution of probability densities in statistics. The second is an extension of the concept of interval arithmetic operating on bandwidths such as shown in Figure 1.

1. Point by Point Computations.

In general, for two numbers, whether in functional form such as A and C in Figure 1, or discrete form such as B, Zadeh's extension principle[2.] can be used to combine these numbers.

$$\mu_{x*y}(w) = \max(\min(\mu(x), \mu(y))$$

where w=u*v and * is any operation. Consider a simple example as shown in Figure 2. Let X and Y be triangular numbers such that

$$X = \underset{\sim}{4} = 0/3 + .1/3.1 + .2/3.2 + .3/3.3 + .4/3.4 + .5/3.5$$
$$+.6/3.6 + .7/3.7 + .8/3.8 + .9/3.9 + 1./4.0 + \cdots$$

in $\mu(X)/X$ notation where + again represents union and

$$Y = \underset{\sim}{12} = 0/11 + .1/11.1 + \ldots + 1./12. + \ldots$$

Let $Z = X+Y$. To calculate this sum, let $Y = Z-X$, then

$$\mu_z(z) = \max(\min(\mu_x(x), \mu_y(z-x)).$$

Performing this calculation for a sample value of $Z = 15$ gives Table 1 and $\mu(15) = \max[0., .1, .2, .3, .4, .5] = .5$. Dubois & Prade (1978) showed that two fuzzy numbers with the same arbitrary shape can be operated on to give a resultant with the same shape - or nearly so. This is true exactly for addition and subtraction and approximately so for multiplication and division. These other operations are given by:

$$\mu_{x-y}(z) = \max(\min(\mu_x(x), \mu_y(z+x))$$

$$\mu_{x.y}(z) = \max(\min(\mu_x(x), \mu_y(z/x))$$

$$\mu_{x/y}(z) = \max(\min(\mu_x(x), \mu_y(x/z)).$$

These operations are analogous to the convolution of probability densities, but with an important difference - in the probabilistic framework, normal densities possess this shape invariance under addition, while under division with normal densities, the resultant is far from normal (a Cauchy density). Intuitively, the reason for this is that in statistics one manipulates densities, or the area under the curve, while in fuzzy arithmetic, one uses the surface of the curve.

There are obvious problems with this method of combining fuzzy numbers. The first is the enormous number of computations involved. Second, if it is assumed that the number has a functional form, as a triangle, for example, then the number of steps one takes in X is crucial. For instance, in our simple case of Z = 15, if only the even values of X are used,

$$\mu(15) = \max(0., .2, .4) = .4$$

which is clearly in error. The next sections of this paper deal with these questions. First an analysis was undertaken to ascertain how many steps are necessary to reach "convergence" and how shape invariant distributions are under multiplication and division. This analysis was done with triangular distributions, although it is equally applicable to any shape as long as the same functional form is used for all the numbers in any sequence of operations.

Computer Analysis of Triangular Distributions

Computer programs were written in Fortran IV to perform point by point multiplication and division operations on triangular numbers. Any shape could have been chosen from the trivial case of a rectangular distribution to a normal distribution with a cutoff outside which $\mu = 0$, to a shape such as in Figure 1,B that is not defined by a function. In order to normalize these fuzzy numbers to some extent, the support was defined in terms of a percentage of the number. The number of steps in X which determined the number of calculations per point in Z was allowed to vary. It can be seen in Figures 3-7 that for a number with a support of $\pm$ 10% of that number

the triangularity remains excellent, providing one is willing to accept differing slopes to the left and right of the peak. However, for $\pm$ 50% of the number, there is considerable deviation. If the triangular invariance is an important consideration, a measure could be devised such that when the deviation of the computed points from the straight line from $\mu = 1$ to $\mu = 0$ exceeded a certain cutoff, shape invariance was no longer assumed.

Another question that arises with this computational method is how many steps must be taken in X to give convergence in Z. Since the computation is a max/min operation, convergence for any point in Z is not a simple iterative procedure. If one happens, with only a few steps, to cycle through the value of X and Y that produces the maximum possible value of min (μ_x, μ_y), then Z has converged for that point and it will remain constant as the number of steps is increased. Hence this analysis consists of a comparison of the maximum deviation in μ at a point Z as the number of steps in X increases from 50 to 1000. The five cases represented in Figures 3-7 were computed. A sample of the output is shown in Table 2 and the results for all five cases in Table 3.

First, as would be expected, within any matrix, as the number of steps are increased, the difference in the maximum value of μ does not change at 1000 steps as compared to 50, 100 and 500. Secondly, for triangular distributions, these maximum differences in μ at specified points in Z are invariant under changing the values of X and Y, the range of their supports, or the operations performed. So one

can conclude that a relatively small number of calculations per point, e.g. 50, is probably adequate.

Intuitively, this invariance can be understood by considering a simple example. Consider two cases where X = 3 but the bandwidth differs as in Figure 8. If n steps are taken in X, the X values are different in X1 and X2 but the μ values are the same. Each value of Z = X-Y or Y = Z/X and hence μ_z and μ_y are constrained by the chosen values of X; the only parameter of concern is the number of steps taken in X to calculate Z.

If computations of this magnitude were necessary, arithmetic operations on fuzzy numbers would not be practically feasible. Another approach has been developed by Dubois and Prade (1978).

2. Parametric Representation and Computation.

Dubois & Prade have represented fuzzy numbers in terms of their peak values with membership at that point being μ =1., and a bandwidth, usually at μ =.5. This is shown in C in Figure 1. It is possible with this representation to simply perform arithmetic operations on fuzzy numbers. It does not matter at what grade of membership the bandwidth is chosen as long as it is the same in all numbers in a given set of calculations. Table 4 gives the results of all 4 operations for all possible values of X and Y. Addition and subtraction are exact, while multiplication and division are approximate. That is, for addition and subtraction, the bandwidth is

the sum of the bandwidths and the peak value is the sum or difference of the peak values. In multiplication and division the peak value is the product or quotient while the normalized bandwidth of the resultant is approximated by the sum of the normalized bandwidths. To get some idea of the magnitude of error, Z at the $\mu = .5$ point was calculated for the cases in Figures 3,4,6,7 with the parametric formulas, and then μ for that point was determined pointwise using 1000 steps in X. The results are shown in Table 5. It can be seen that for a bandwidth of .05 of the peak, the parametric representation provides a near perfect calculation for multiplication, in that the triangular shape is preserved. Since division skews the resultant distribution more than multiplication, the parametric approximation for division does not produce such accurate results. However, it should be possible to derive a better approximation, since the error seems to be constant at Z as determined by the parametric formulae for $\mu = .5$.

3. Parametric Operations on Algebraic Equations

Assume for the moment that one is willing to tolerate the error caused by parametric operations, or is willing to perform a point by point operation for division and multiplication when the bandwidth exceeds an arbitrarily chosen percentage of the peak. How then do these fuzzy numbers behave after a sequence of operations? A Fortran IV program was written to perform any sequence of operations on any number of variables. As an example, let the operation be

$$R = E + [F * C * B] - [D / A] + F$$

Let the variables be "reasonably" defined so that the value of the bandwidth at $\mu = .5$ is equal to 10% of the peak value. Table 6 shows this operation.

If the bandwidth at $\mu = .5$ is allowed to be 50% of the peak for each number A to F, then

$$R = [\ 26.938 \qquad -33.250 \qquad 64.813]$$

Interpretation

Interpretation of the degree of fuzziness of the resultant is not an obvious matter. When two symmetric positive numbers are added, say $A + C = 9.$, the resulting bandwidths, .9, are still 10% of the resultant. However, for $A - C = -1.$, the bandwidths are again .9. Hence it is not possible to speak in terms of the bandwidths being a percentage of the original variable. Are the results of these operations

$$A + C = [.9 \quad 9. \quad .9]$$

and

$$A - C = [.9 \quad -1. \quad .9]$$

really of different degrees of fuzziness given the original variables? This problem is intensified for multiplication and division, and if one were to add several large only slightly fuzzy numbers at the end of a series of operations, the resultant might appear quite good, indeed.

Our perception of the numbers, then has to do with the non-normalized bandwidth, at least in the range of zero. Although this technique of characterizing a number as a percentage of the bandwidth is useful as a description of the original variables, when the result of an operation is considered it is as if the peak, in some sense, becomes unimportant in this region of the real number line. Possibly one should revert, at least conceptually to a representation of the resultant suggested by interval analysis. For instance,

$$A + C = [8.1, 9.9] \qquad \text{peak} = 9.0$$

$$A - C = [-1.9, 0.1] \qquad \text{peak} = -1.0$$

emphasizes that it is the magnitude of the interval itself that is the most important datum to result from the computation. As the peak value gets large this psychological perception of the size of the bandwidth again could be interpreted as a percentage of the peak. For example, let

$$A = [80, 100] \quad \text{or } [10, 90, 10]$$

$$B = [800, 1000] \quad \text{or } [100, 900, 100]$$

and both give approximately the same sense of precision.

Another facet of the difficulty of the interpretation of interval numbers or fuzzy numbers arises from the non-distributivity of these numbers. Different intervals are obtained for Y if, for example, one performs $Y = X/(X-1)$ and $Y = 1 + 1/(X-1)$. For example, let X be a fuzzy number $X = (1,5,2)$ and 1 a constant $1 = (0,1,0)$. Then

$$Y = X/[X-1] = [.75,\ 1.25,\ .9375]$$

$$Y = 1 + 1/[X-1] = [.125,\ 1.25,\ .0625].$$

The smaller bandwidth in the second expression is due to the multiple use of the constant 1. However in the example computed above,

$$R = E + [F*C*B] - [D/A] + F$$

and

$$R = E + F[C*B+1] - D/A$$

give the same resultant. If sequences of operations are to be performed utilizing any variables more than once, this property of non-distributivity must be taken into account. Moore (1966) discusses a scheme for selecting the expression to give the minimum interval, but it is complex.

It should be noted that these techniques are related in a broad sense to traditional error analysis that is concerned with the propagation of errors in numerical calculations. An example of such a system is process graphs (e.g. Dorn & McCracken, 1972) which use different formulae than those used here but hit some of the same philosophical problems with regard to absolute vs. relative errors. It seems worth considering whether the types of errors traditionally considered in numerical analysis could be viewed in a fuzzy framework. Numerical analysis tends to be concerned with inherent errors in the data, truncation errors and roundoff errors. Only the first is of interest here. For an instrument reading or a time interval, if only a few significant figures are given, as in, say 1.5 seconds, we may be able to place bounds on the numbers, for instance, 1.5 ± 0.2 seconds. If these numbers could be philosophically justified as fuzzy numbers, then fuzzy arithmetic could be applied to physical measurements or distance measurements as well as questionnaire data within the same

analysis. This could lead to a powerful technique for data validation, for instance.[3]

Reciprocally, the step by step computation in Table 6 might be viewed as a kind of process graph where each operation is a node. This formulation is useful if one wishes to see how the numbers spread under successive operations, particularly in the case of alternative algebraic representations of the same quantity.

Some Further Thoughts on Fuzzy Numbers.

These analyses suggest two interesting extensions. The first concerns manipulating different shapes of numbers while the second relates to possible applications of the theory of fuzzy numbers.

1. Numbers with different shapes. Suppose that the shape of the distributions were not triangular, but of another form. Zadeh (1973) has suggested using a piecewise quadratic function, and this seems a likely shape to investigate since it more closely resembles a normal distribution than does a triangular distribution. The difficulty in using a normal distribution is that its support is infinite, although an arbitrary cutoff could be made, beyond which it is considered to have $\mu = 0$. Another form that frequently occurs in real-world applications is that of a rectangular distribution. This resembles an interval number where $\mu = 1$ or $\mu =$ constant for every value within the support. Two approaches in the literature for dealing with these varied forms of numbers seem immediately relevant. First, Dubois and

Prade (1978) combine distributions in a piecewise fashion so more than one functional form could be used in the same operation. Second, Jain (1976) considers rectangular numbers alone within the context of tolerance analysis in network design where components and variables can take a range of values. By using a different definition for the sum of two fuzzy variables, rather than Zadeh's formulation he shows that two sets that have uniform μ within each set add such that the values near the center of the sum have higher grades of membership than those at the extremes of the resultant distribution.

This points up an interesting facet in the development of fuzzy set theory. Zadeh's initial formulation arose from the semantics of applied systems analysis while other, more formal developments, have emanated from many varients of multivalued logic (e.g. Gaines, 1976) or by analogy with probability theory (e.g. Nahmias, 1978). These different systems result in different formulations of fuzzy variables and different operations on them. Which is "correct" is addressed by Gaines (1976):

> ... only the semantics of particular applications can determine which is the appropriate choice. In many practical situations these semantics are so fuzzy that the choice does not matter over a wide range of possibilities. Developments of the mathematical foundations of vague reasoning need to take this into account and not attempt to introduce a new level of arbitrary precision in the metalanguage - detailed and specific arguements as to what are the "right" functions seem singularly inappropriate to the subject area. We need integrative, broadly based theories with strong intuitive appeal.

2. Discrimination of fuzzy numbers. This work was initially motivated by the idea of using fuzzy numbers in data analysis, for example in data validation where two combinations of different variables should give the same result. But if the resultants are A = (.5,3.,.5) and B = (1.,4.,1.), one must have a way of measuring to what degree the numbers agree. Complete overlap would give $\mu = 1.$, while for complete separation $\mu = 0$. This type of measure relates to that developed by Enta (1978) who describes a technique for assessing the degree of separation among possibility distributions. This topic will be the subject of a future paper.

Footnotes

1. A normal density with mean 0, SD 1 has a maximum Y ordinate of .3984, hence at -1SD, Y = .2420, and μ [X=-1] = .3984 X .2420 = .6067.

2. In probability theory, the extension principle is analogous to the expression for the probability density induced by a mapping. See, e.g. Thomasian (1969).

3. An example of the use of fuzzy set theory in the analysis of physical quantities is given by Jain (1976). In this paper he discusses a method for the analysis of the effects of subjective and objective tolerances in networks design. Since components used in practically every type of networks and systems may vary over a range of values, tolerance analysis is undertaken to show the effects of these variations. Sometimes statistical distributions of component values are available and standard techniques can be effectively used. At other times, however, the component variable values and their associated uncertainties are given linguistically as "small" or "large" or "about 100 KΩ". Jain suggests that fuzzy set theory could be used here, as well as in cases that could be conventionally analysed, but in which one wishes to consider the distribution of published component values as fuzzy.

References

Dubois, D. & Prade, H., Operations on Fuzzy Numbers. International Journal of Systems Sciences, 1978, to be published.

Dorn W. S. & McCracken, D. D., Numerical Methods with Fortran IV Case Studies, New York: John Wiley & Sons, Inc., 1972.

Enta, Y. Discriminative effect of information: The third measure information. No. UCB/ERL M78/62 University of California, Berkeley, May, 1978.

Gaines, B.R. Foundations of fuzzy reasoning. International Journal of Man-Machine Studies, 1976, 6, 623-668.

Jain, R. Tolerance Analysis using fuzzy sets. International Journal of Systems Science, 1976, 7, [12], 1393-1401.

Moore, R.E., Interval Analysis. Englewood Cliffs, N.J.: Prentice Hall, Inc., 1969

Nahmias, S., Fuzzy Variables, Fuzzy Sets and Systems 1978, 1, 97-110

Mizumoto, M. & Tanaka, K., Algebraic Properties of Fuzzy Numbers, Proceedings of the International Conference on Cybernetics and Society Washington, D.C. 1976, 559-564.

Zadeh, L. A., Fuzzy Sets, Information and Control, 1965, 8, 338-353.

Zadeh, L. A., The Concept of a Linguistic Variable and its Applications to approximate reasoning. Part I, Information Science, 1975, 8, 199-249; Part II 8, 301-357; Part III 9, 43-80.

Zadeh, L. A., Fuzzy Sets as a Basis for a Theory of Possibility. Fuzzy Sets and Systems, 1978, 1, 3-28.

APPENDIX 1

Derivation of formulae in Table 4

These formulae were derived by Dubois and Prade (1978). A few are explained here for ease in understanding Table 4.

Basically there are two rules:

Addition $X \oplus Y$

Multiplication $X \oplus Y$

these generate

Subtraction $X \oplus (-Y)$

Division $X \otimes (1/Y)$

Dubois and Prade show that for

$$Y = (\gamma, Y, \delta)$$
$$1/Y = (\delta/Y^2, 1/Y, \gamma/Y^2)$$

Using the parametric forms, without the second order correction factor for simplicity:

For $X > 0 \quad Y > 0$

$$(\alpha, X, \beta) \oplus (\gamma, Y, \delta) = (\alpha + \gamma, X + Y, \beta + \delta)$$
$$(\alpha, X, \beta) \otimes (\gamma, Y, \delta) = (X\alpha + Y\alpha, XY, X\delta + Y\beta)$$

Subtraction and Division can be derived from these:

$$\begin{aligned}(\alpha, X, \beta) \ominus (\gamma, Y, \delta) &= (\alpha, X, \beta) \oplus (-(\gamma, Y, \delta)) \\ &= (\alpha, X, \beta) \oplus (\delta, -Y, \gamma) \\ &= (\alpha + \delta, X - Y, \beta + \gamma)\end{aligned}$$

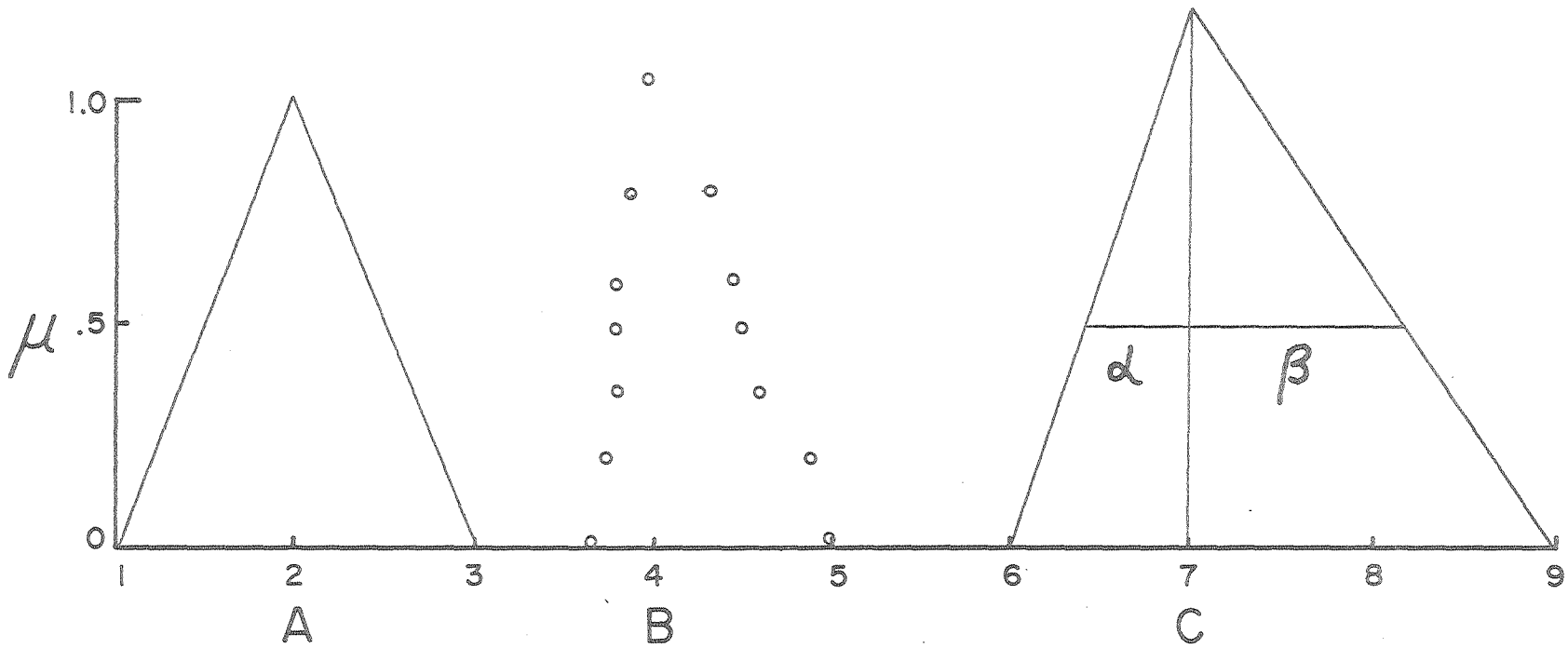
$$\begin{aligned}(\alpha, X, \beta) \oslash (\gamma, Y, \delta) &= (\alpha, X, \beta) \otimes (\delta/Y^2, 1/Y^2, \gamma/Y^2) \\ &= (X\delta/Y^2 + \alpha/Y, X/Y, X\gamma/Y^2 + \beta/Y) \\ &= ((X\delta + Y\alpha)/Y^2, X/Y, (X\gamma + Y\beta)/Y^2)\end{aligned}$$

For $X < 0 \quad Y \geqslant 0$

$$\begin{aligned}(-(\alpha, X, \beta)) \oplus (\gamma, Y, \delta) &= (\beta, -X, \alpha) \oplus (\gamma, Y, \delta) \\ &= (\beta + \gamma, X + Y, \alpha + \delta)\end{aligned}$$

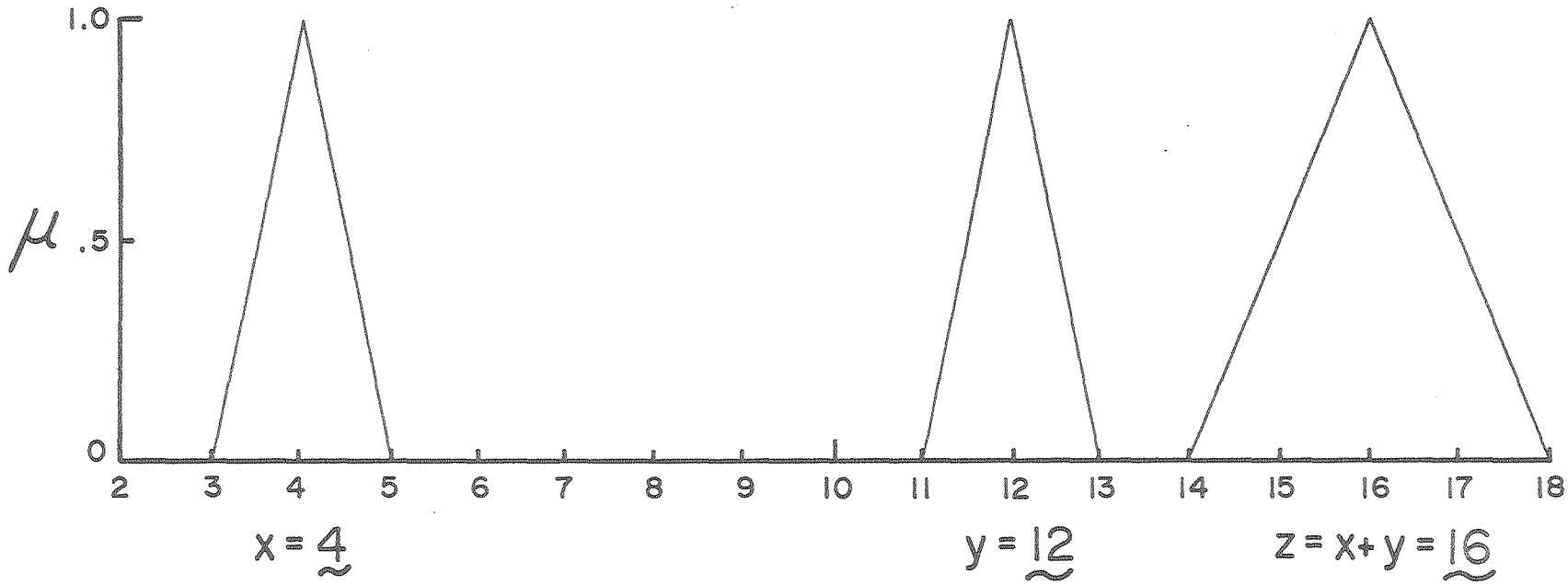
$$\begin{aligned}-[(-(\alpha, X, \beta)) \otimes (\gamma, Y, \delta)] &= -[(\beta, -X, \alpha) \otimes (\gamma, Y, \delta)] \\ &= -[-X\gamma + Y\beta, XY, -X\delta + Y\alpha] \\ &= (-X\delta + Y\alpha, XY, -X\gamma + Y\beta)\end{aligned}$$

All the other formulae can be derived following these procedures.



XBL 793-8781

FIGURE 1: 2, 4 and 7



XBL 793-8782

FIGURE 2: Fuzzy Addition - $\underset{\sim}{4} + \underset{\sim}{12} = \underset{\sim}{16}$

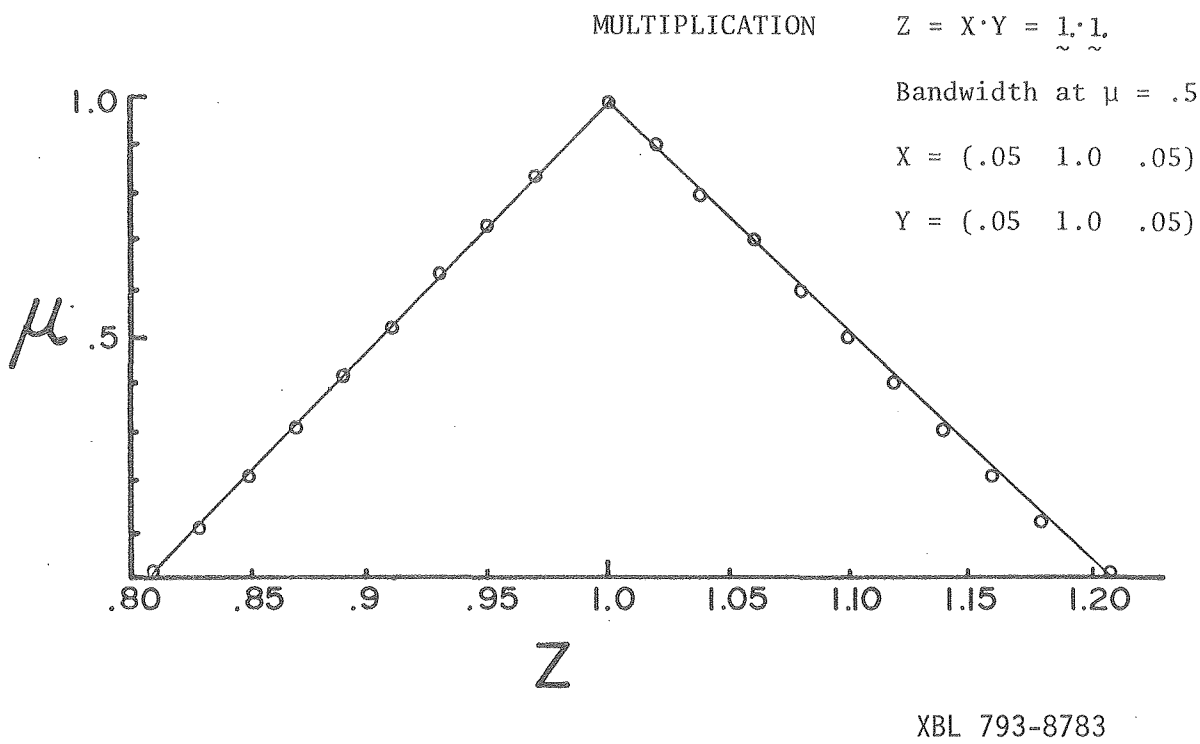


FIGURE 3: Resultant of 1.·1.; 1000 calculations per point.

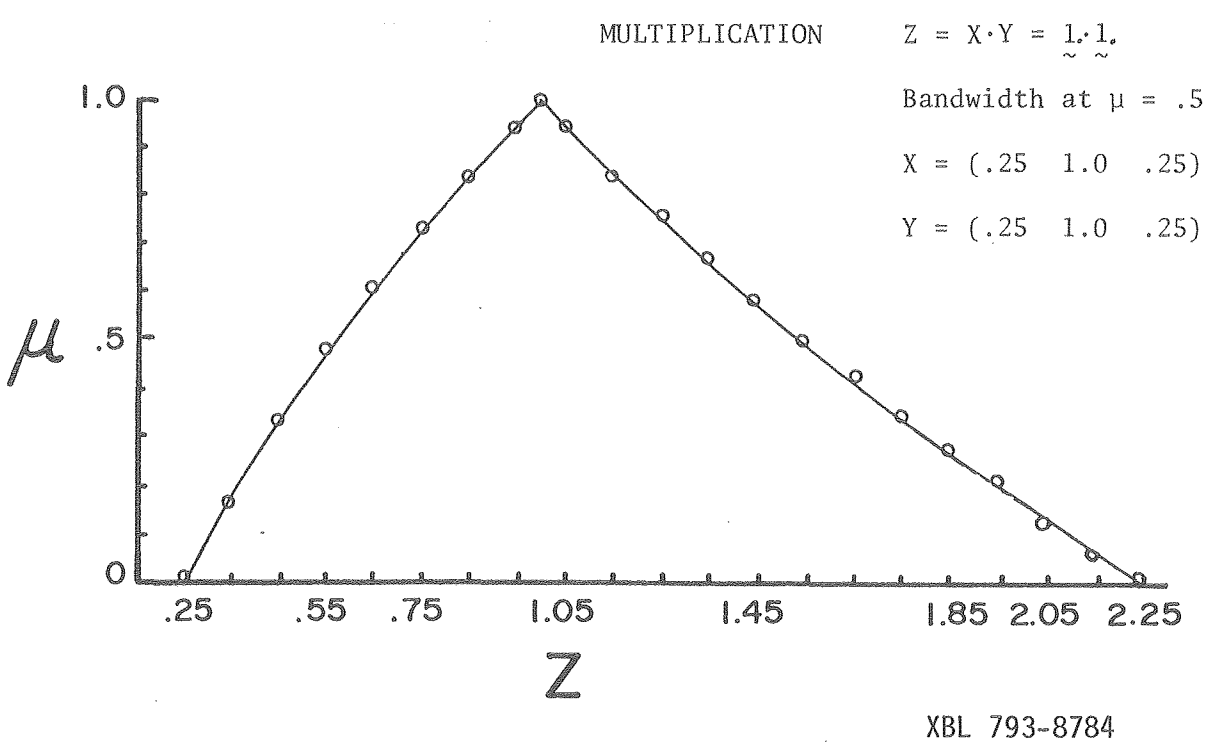
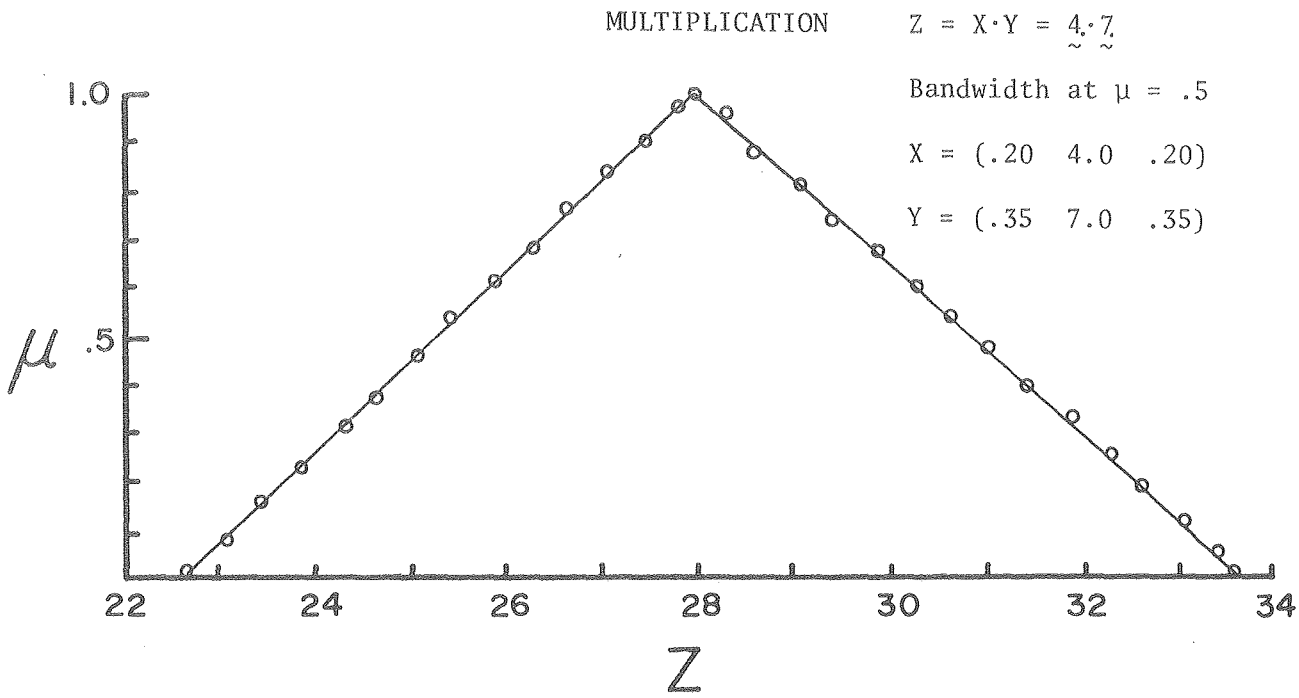
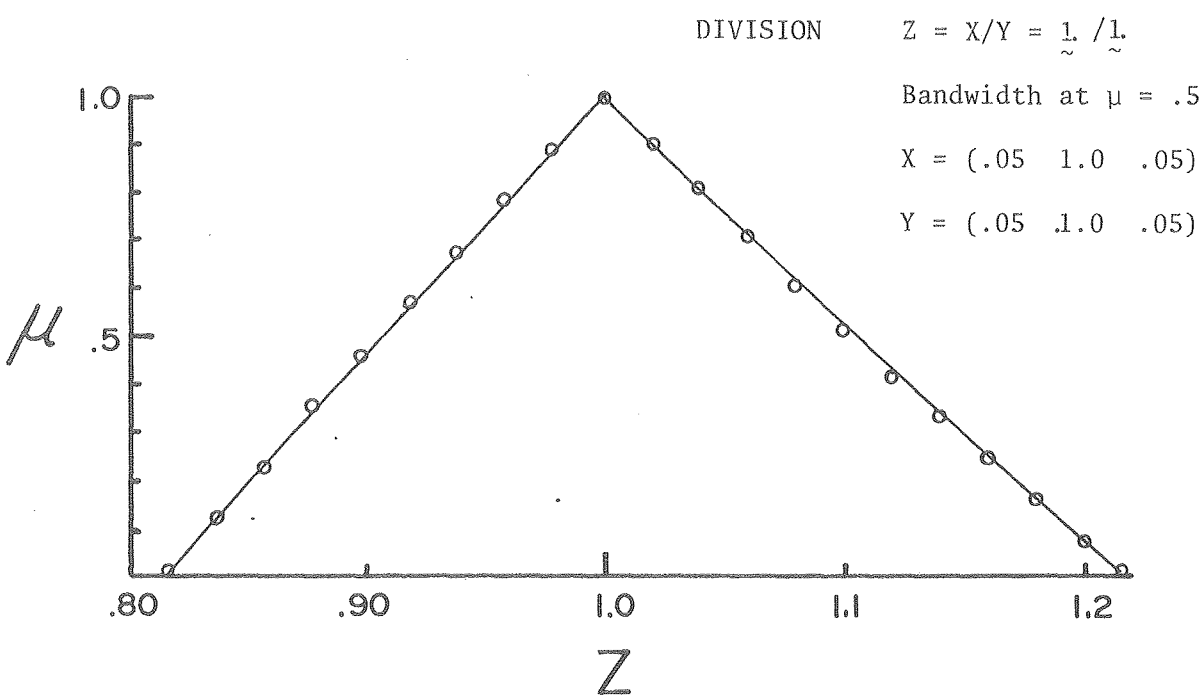


FIGURE 4: Resultant of 1.·1.; 1000 calculations per point.



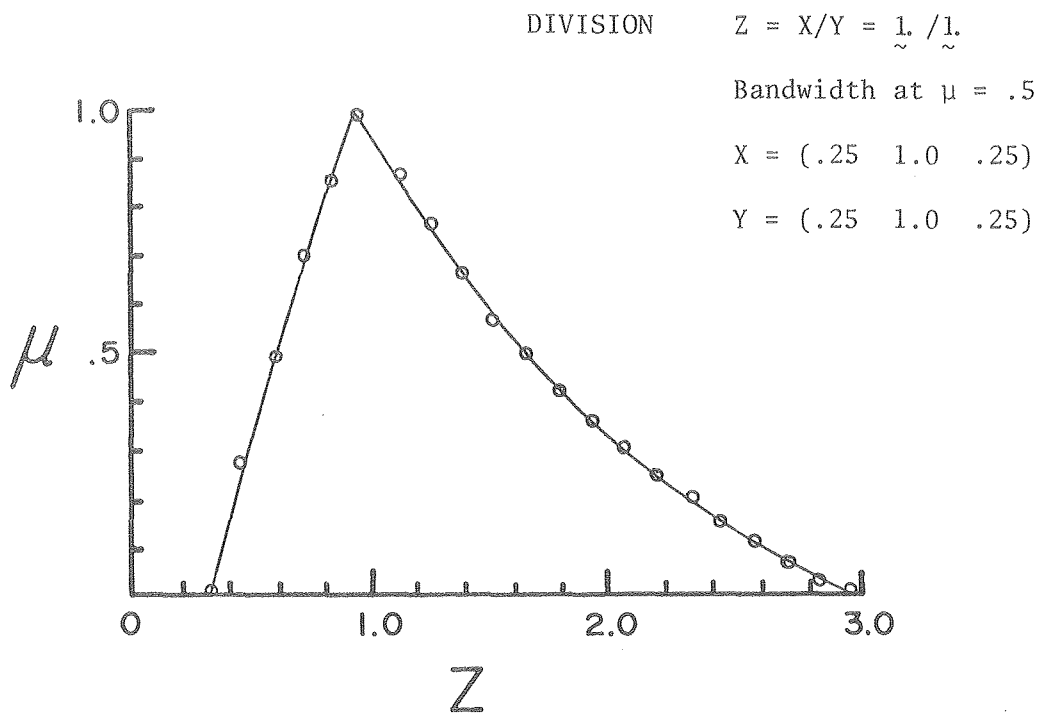
XBL 793-8785

FIGURE 5: Resultant of 4.·7.; 1000 calculations per point.



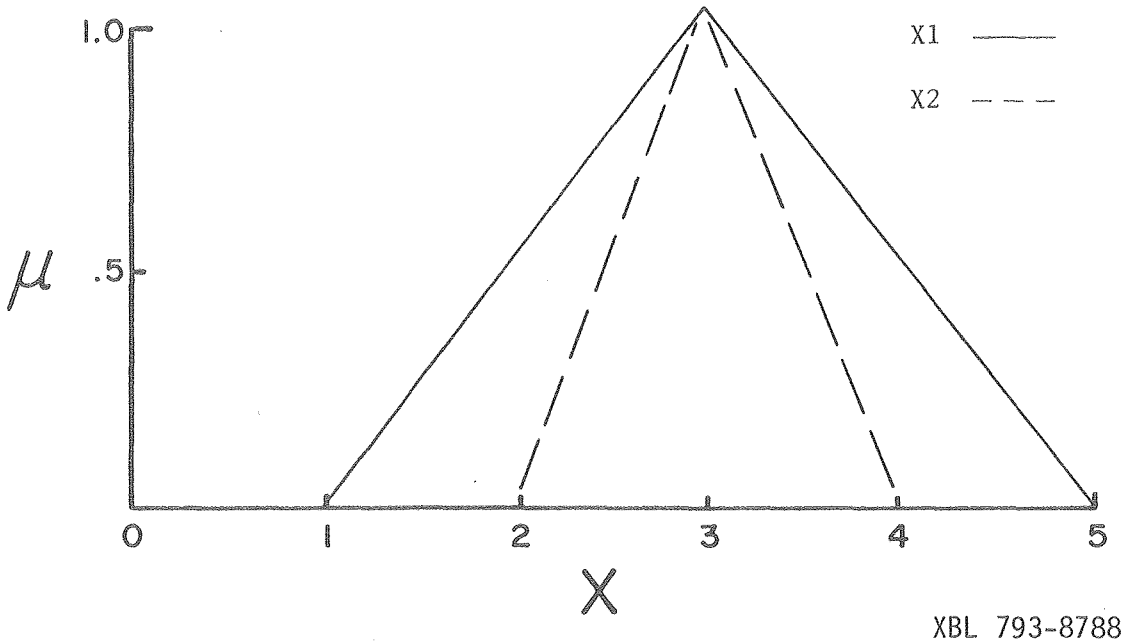
XBL 793-8786

FIGURE 6: Resultant of 1. /1. ; 1000 calculations per point.



XBL 793-8787

FIGURE 7: Resultant of 1. /1. ; 1000 calculations per point.



XBL 793-8788

FIGURE 8: X = 3 with different bandwidth.

Z	X	Y=Z-X	μ(X)	μ(Y)	min(μ(X),μ(Y))
15	3.	12.	0.	1.	0.
	3.1	11.9	.1	.9	.1
	3.2	11.8	.2	.8	.2
	3.3	11.7	.3	.7	.3
	3.4	11.6	.4	.6	.4
	3.5	11.5	.5	.5	.5
	.				
	.				
	.				
	5.0	10.0	0.	0.	0.

TABLE 1: μ for Z = 15

		Calculation per point			
		50	100	500	1000
	Z=X/Y	$\mu_{Z\ 50}$	$\mu_{Z\ 100}$	$\mu_{Z\ 500}$	$\mu_{Z\ 1000}$
1	.333333	.000000	.000000	.000000	.000000
2	.466666	.257141	.260000	.272000	.272000
3	.600000	.480000	.499999	.499999	.499999
4	.733333	.680000	.681818	.692000	.692000
5	.866667	.840000	.853846	.856000	.856154
6	1.000000	1.000000	1.000000	1.000000	1.000000
7	1.133333	.870588	.870588	.874118	.874118
8	1.266667	.760000	.760000	.764000	.764000
9	1.400000	.657143	.660000	.665714	.666000
10	1.533334	.565217	.578261	.578261	.578261
11	1.666667	.488000	.500000	.500000	.500000
12	1.800000	.422222	.422222	.428000	.428000
13	1.933334	.360000	.360000	.363448	.363448
14	2.066667	.296774	.300000	.304000	.304000
15	2.200000	.240000	.245454	.249091	.250000
16	2.333334	.200000	.200000	.200000	.200000
17	2.466667	.151351	.151351	.152973	.153784
18	2.600000	.107692	.107692	.110769	.110769
19	2.733333	.068293	.068293	.071219	.071219
20	2.866667	.032558	.032558	.033953	.034000
21	3.000000	.000000	.000000	.000000	.000000

TABLE 2: DIVISION Z = X/Y

for X = (.25 1. .25)

Y = (.25 1. .25)

Z = (.4275 1. .5625)

MULTIPLICATION 1. x 1.

Range of X = .9 - 1.1
Range of Y = .9 - 1.1
Range of Z = .81 - 1.21
(±10% of X and Y)

steps in X	50	100	500	1000
50	-	.020	.019	.019
100		-	.008	.008
500			-	.002
1000				-

X = .5 - 1.5
Y = .5 - 1.5
Z = .25 - 2.25
(±50% of X and Y)

steps in X	50	100	500	1000
50	-	.020	.020	.020
100		-	.008	.009
500			-	.002
1000				-

All entries are the maximum value in the difference between μ_Z columns as shown in the sample in Table 2.

MULTIPLICATION 4. x 7.

Range of X = 3.60 - 4.40
Range of Y = 6.30 - 7.70
Range of Z = 22.68 - 33.88
(±10% of X and Y)

steps in X	50	100	500	1000
50	-	.019	.019	.019
100		-	.008	.008
500			-	.002
1000				-

DIVISION 1./ 1.

Range of X = .9 - 1.1
Range of Y = .9 - 1.1
Range of Z = .8182 - 1.222
(±10% of X and Y)

steps in X	50	100	500	1000
50	-	.019	.018	.018
100		-	.007	.007
500			-	.002
1000				-

X = .5 - 1.5
Y = .5 - 1.5
Z = .3333 - 3.0
(±50% of X and Y)

steps in X	50	100	500	1000
50	-	.019	.019	.019
100		-	.012	.012
500			-	.001
1000				-

TABLE 3: Matrices Showing the Maximum Value of the Difference in μ_Z as the Number of Points in X Changes.

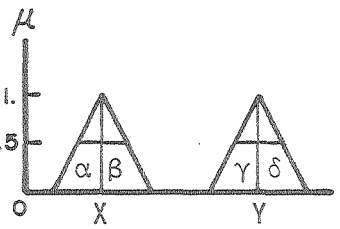
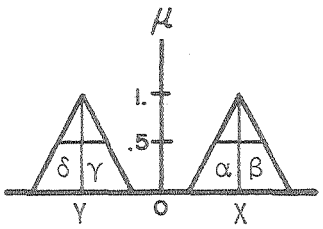
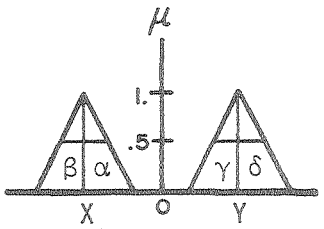
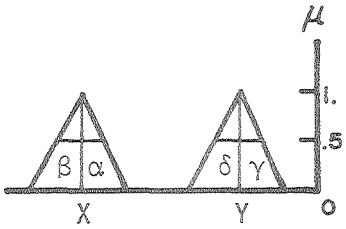
OPERATION	BAND (LEFT OR RIGHT)	 X ⩾ 0 Y ⩾ 0	 X ⩾ 0 Y < 0	 X < 0 Y ⩾ 0	 X < 0 Y < 0
X/Y	L	$(X\delta + Y\alpha - \alpha\delta)/Y^2$	$(X\delta - Y\beta + \beta\delta)/Y^2$	$(-X\gamma + Y\alpha - \alpha\gamma)/Y^2$	$(-X\gamma - Y\beta - \beta\gamma)/Y^2$
	R	$(X\gamma + Y\beta + \gamma\beta)/Y^2$	$(X\gamma - Y\alpha - \alpha\gamma)/Y^2$	$(-X\delta + Y\beta + \beta\delta)/Y^2$	$(-X\delta - Y\alpha + \alpha\delta)/Y^2$
X*Y	L	$X\gamma + Y\alpha - \alpha\gamma$	$X\gamma - Y\beta + \beta\gamma$	$-X\delta + Y\alpha + \alpha\delta$	$-X\delta - Y\beta - \beta\delta$
	R	$X\delta + Y\beta + \beta\delta$	$X\delta - Y\alpha - \alpha\delta$	$-X\gamma + Y\beta - \beta\gamma$	$-X\gamma - Y\alpha + \alpha\gamma$
X-Y	L	$\alpha + \delta$	$\alpha + \gamma$	$\beta + \delta$	$\beta + \gamma$
	R	$\beta + \gamma$	$\beta + \delta$	$\alpha + \gamma$	$\alpha + \delta$
X+Y	L	$\alpha + \gamma$	$\alpha + \delta$	$\beta + \gamma$	$\beta + \delta$
	R	$\beta + \delta$	$\beta + \gamma$	$\alpha + \delta$	$\alpha + \gamma$

TABLE 4: ARITHMETIC OPERATIONS ON PARAMETERIZED FUZZY NUMBERS (see APPENDIX 1)

Division Z=X/Y = 1./1.

BANDWIDTH	Z		μ_Z (1000 steps in X)	μ_Z (parametrically)
± 5% mean	left	.9025	.487	.500
	right	1.1025	.512	.500
± 25% mean	left	.5625	.440	.500
	right	1.5625	.560	.500
Division Z=X/Y = 4./7.				
± 5% mean	left	.5157	.487	.500
	right	.6300	.512	.500
± 25% mean	left	.3214	.440	.500
	right	.8929	.560	.550
Multiplication Z = X*Y = 1.*1.				
± 5% mean	left	.9025	.500	.500
	right	1.1025	.500	.500
± 25% mean	left	.5625	.500	.500
	right	1.5625	.500	.500
Multiplication Z= X*Y = 4.*7.				
± 5% mean	left	25.27	.500	.500
	right	30.87	.500	.500
± 25% mean	left	15.75	.500	.500
	right	43.75	.500	.500

TABLE 5: μ_Z Computed Point by Point and Parametrically

VARIABLE	LEFT μ =.5 POINT	MEAN	RIGHT μ = .5 POINT
A	.40	4.0	.40
B	.10	1.0	.10
C	.50	5.0	.50
D	.30	-3.0	.30
E	.40	-4.0	.40
F	.50	-5.0	.50

Then step by step computations give:

M=D/A	.158	-.750	.143
N=F*C	4.750	-25.0	5.250
O=N*B	6.775	-25.0	8.275
P=O-M	6.918	-24.250	8.433
Q=E+P	7.318	-28.250	8.833
R=Q+F	7.818	-33.250	9.333

TABLE 6: R = E + [F*C*B] - [D/A] + F

This report was done with support from the Department of Energy. Any conclusions or opinions expressed in this report represent solely those of the author(s) and not necessarily those of The Regents of the University of California, the Lawrence Berkeley Laboratory or the Department of Energy.

Reference to a company or product name does not imply approval or recommendation of the product by the University of California or the U.S. Department of Energy to the exclusion of others that may be suitable.

TECHNICAL INFORMATION DEPARTMENT
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720

AAX968
LBL Libraries