Conductivity Porosity Relationships in Rocks

T. R. Madden

The geometry of the interconnecting pores and cracks of a rock has a strong effect on the electrical properties of water saturated rocks. Although a good deal of empirical data has been collected, the theory governing the geometrical effects is poorly developed. At low ionic concentrations other factors besides the geometry of the rock interstices become important. These factors will be considered later, but here we wish to examine the geometric factors and we shall consider the pore fluids as highly concentrated. In such cases the rock resistivity becomes proportional to the fluid resistivity, and the proportionality factor carries the geometric information.

$$\rho_{rock} = I\ F\ \rho_{solution} \tag{1.1}$$

F = formation factor

I = resistivity index

F is controlled by the geometry of the interstices and I is controlled by the geometry of the fluid within the interstices. For water saturated rocks I is equal to one and can be left out of equation (1.1).

If we define Por = porosity

S = fraction of water saturation

it is found that

$$F \approx a\ Por^{-m}\ ,\ a \sim 1 \tag{1.2}$$

$$I \approx S^{-n} \qquad m \approx n \approx +2$$

This power law dependence of resistivity on porosity is known as Archie's Law. It is not an exact relationship, but it does seem to hold approximately over a very wide range of porosities. In its simplest and original form Archie's Law is $\rho_{rock} = \rho_{solution}\ Por^{-m}$ (1.3)

Figure 1 shows data covering three decades of porosity which follows this relationship with $m = 2$.

Another term used to describe a geometric feature of the conduction paths is the tortuosity, T. If the conduction paths had constant cross sectional areas, but their lengths were T_* linear dimensions we would have

$$Por = AT \tag{1.3}$$

A = cross sectional area of pores per unit area

T = length of pores per unit length of sample

This same model would also give

$$\begin{aligned}\sigma_{rock} &= \sigma_{solution}* A/T \\ &= \sigma_{solution}* Por/T^2\end{aligned} \tag{1.4}$$

Archie's Law $\sigma = \sigma_{sol}\ P^2$ would imply that the tortuosity is inversely proportional to the square root of the porosity. For very tight rocks the numerical value one would assign to the tortuosity is too large to have much meaning as an excess length. In such cases it might be more reasonable to consider the tortuosity as indicative of excess pore volume as shown in figure 2. If in such a model we define

T^2 = total pore volume/useful pore volume

we would still retain the same form for equation (1.4).

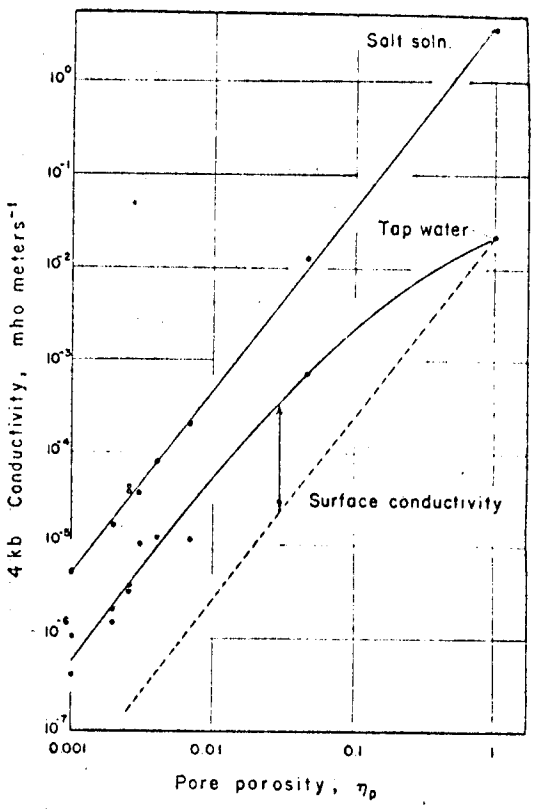


Fig. 1 Conductivity at 4 kb as a function of pore porosity in tap water and in salt solution.

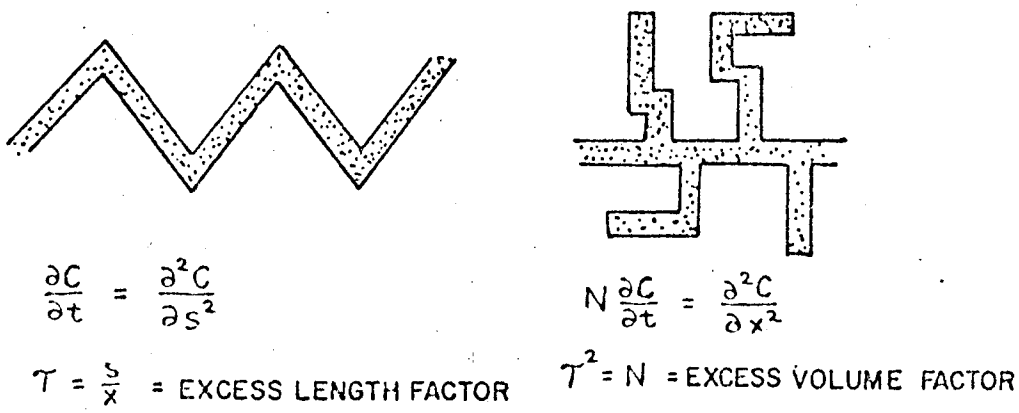


Fig. 2. Interpretations of Tortuosity.

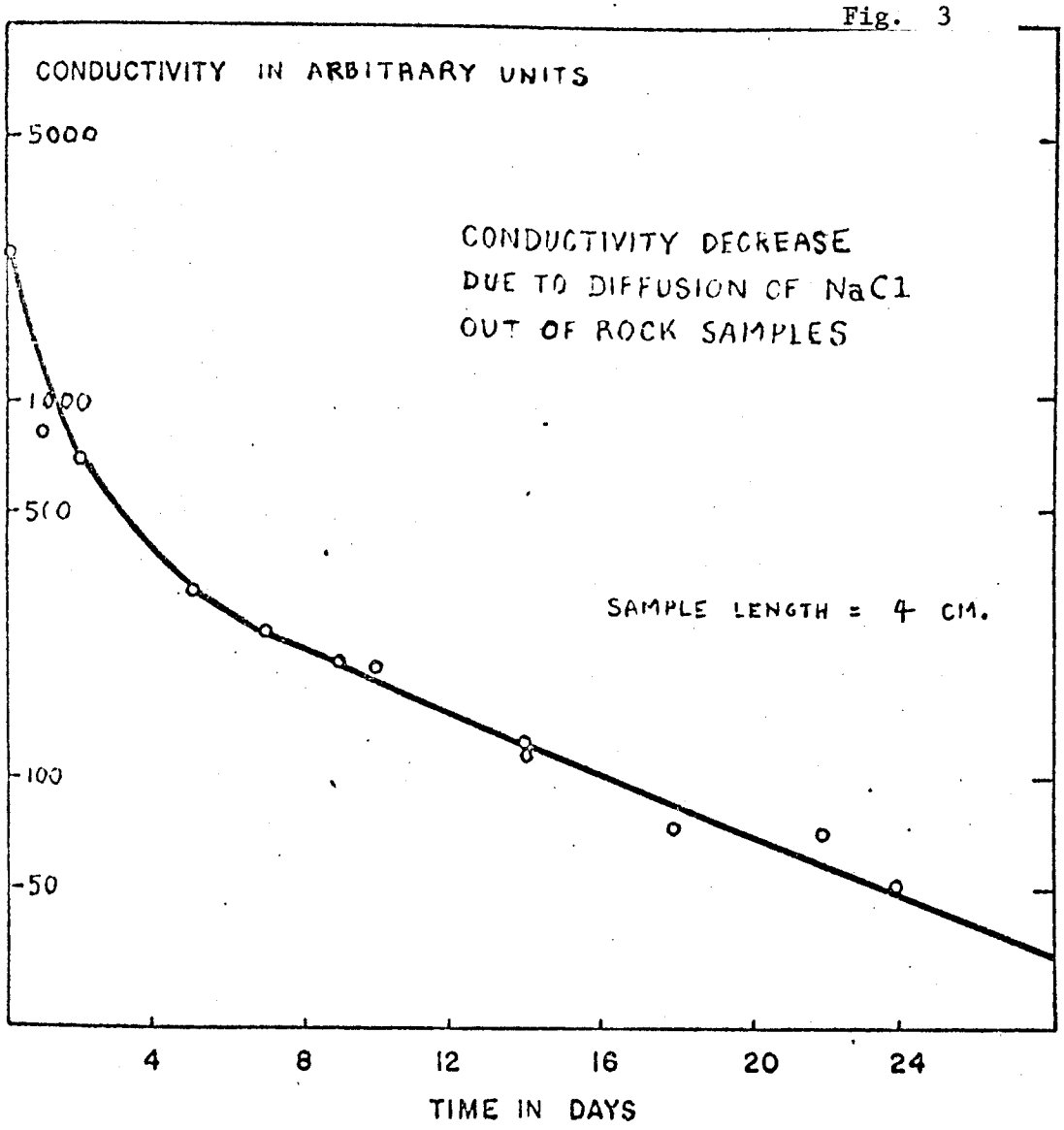
Diffusion time is strongly dependent on distance, and one can indirectly measure tortuosity by timing the diffusion of salt into or out of rock samples, the state of salinity of the fluid within the rock being monitored by the rock's electrical resistance. Figure 3 is an example of such a measurement. If we use the first model of tortuosity representing an excess length factor then the diffusion time will be proportional to $(TL)^2$. The same result is arrived at if we use the second model, since the diffusion flow must fill up a total volume of $T^2 *$ useful volume, and again the diffusion time will be proportional to T^2L^2.

None of these concepts seem very useful in predicting Archie's Law. Figure 4 shows a simple model that does show such a behavior, however. The porosity of each segment is proportional to A, the cross sectional area of the conduction paths. When the two sections are abutted together the common area which will support conduction across the two sections is proportional to Por^2 provided the location of the pore ends are uncorrelated between the two segments. Adding a third section will change this result, but the model is too crude to be taken very seriously.

The wide range of applicability of Archie's Law is intriguing and one cannot help but wonder what principle is involved.

A few years ago in discussing Archie's Law with Bill Brace and Joe Walsh, we considered the possibility of modelling rock conduction with a resistance network. Resistors in place would represent pore regions and missing resistors would represent rock matrix regions. Analogue models were used in a study by Richard Greenberg, where resistors were systematically removed from the network and the changes of the network resistivity were recorded at each step. The choice of resistors to be removed was made randomly.

Fig. 3



$$\sigma \cong \sigma_\infty + \sum B_i e^{-D4\pi^2 t / L_i^2}$$

σ_∞ = final conductivity

B_i = excess conductivity due to saline solution in pore system i

L_i = effective length of i^{th} pore system

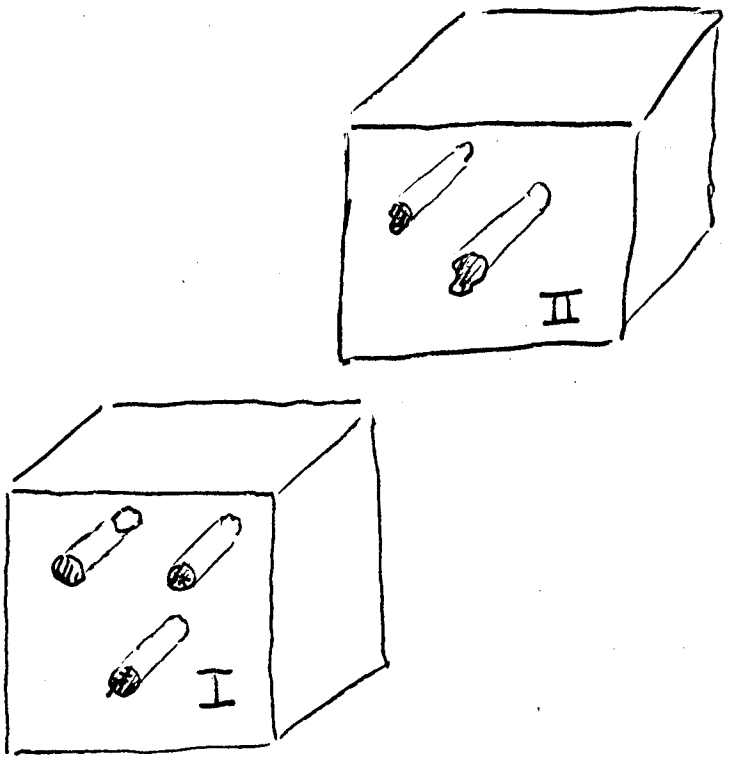


Fig. 4. Model of Archie's Law behavior
If zone I is abutted to zone II and location of pores in I is uncorrelated to pore locations in II
$\sigma_{I\&II} \sim Por^2$

Figure 5 shows one of his results on a cubical network. An Archie type law is observed, but the network suddenly open circuited while a fair fraction of the resistors was still present in the model. At first one might think this open circuiting would be avoided if a larger network had been used, but it turns out this is not the case, and Greenberg's results are not very different from what would be expected with an infinite sized network of the same shape and with the same geometry of interconnections.

There exists a field of study called percolation theory which is concerned with flows in random media. One of the principal results of this theory is the existence of a critical probability for networks. If network elements are randomly missing, the network will open circuit when the probability of an element being in place drops below the critical probability (in the limit of an infinitely sized network). The value of the critical probability depends on the network topology. Table I presents results obtained from Monte Carlo experiments (Vyssotsky et al 1961)

Table I Critical Probabilities

Lattice type	P_c
Triangular	.341 ± .011
Square	.493 ± .013
Cubic	.254 ± .013
Tetrahedral	.390 ± .011
Face Centered Cubic	.125 ± .005

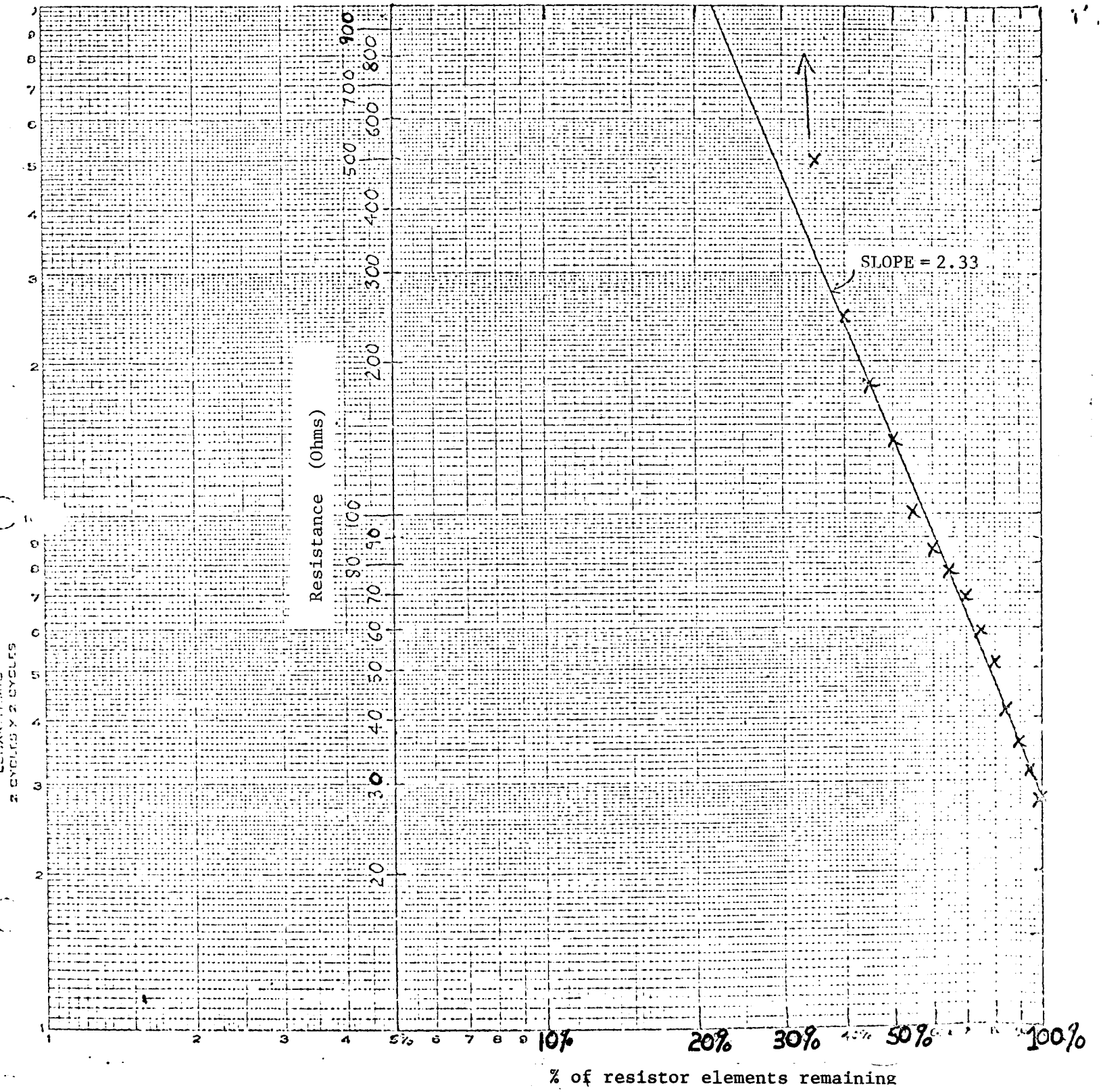


Fig. 5. Results for 3-D Resistor Network with 2000 elements.

The model we were using, although it appeared to follow Archie's Law, could never work to describe low porosity rocks, since they would open circuit. This never seems to happen to actual rocks. The missing factor in the model is the lack of any correlation between the probability of colinear elements being in place. In rocks the cracks and pores that represent the conduction paths have two or one dimension very much longer than the other dimension or dimensions, and this feature is missing from the model. In fact we can go to the extreme case of saying all the pores and cracks are interconnected and this will probably be a much better model. From mechanical compliance measurements one is aware of the fact that cracks are thousands of times longer than they are thick, and if such cracks have a total volume of 0.1% or better it is very difficult to arrange them without their intersecting. If porosity is defined or measured by the volume of liquid that can be put into a rock, then of course all the pore spaces involved must be interconnected. Also, we must remember, if there is not good interconnectivity the rocks would open circuit. For instance in the square networks the probability of an element being isolated is only .015 at the critical probability when the whole network open circuits. For a cubic network this probability is .05 .

On this basis we can make a new network model which involves only the conduction paths of the rock. In this model the network elements are all in place, but the values of the elements are random and represent the distribution of values of conduction in small regions due to variations in crack and pore parameters and statistics. The behavior of these models will depend on the element value distribution functions and on the network topology. The element value distribution function would seem to reflect certain basic statistical properties of the crack

and pore populations, but the network topology seems somewhat arbitrary. It is hoped therefore that the topology does not have a strong influence on the results.

We have looked at a few examples using these ideas and the results have been encouraging. Figure 6 shows some random network calculations using cubical networks. The distribution functions used were all very smooth and covered a wide range of values. Very similar results were obtained with square networks. Each network element is assumed to represent a small enough segment of a rock conduction path that it has essentially no tortuosity and its contribution to the porosity is therefore linearly dependent on its conductivity.

In order to progress further on this problem one needs a more analytic approach to determining the random network properties. The transmission matrix which relates the voltages and currents at one end to the voltages and currents of the other end can be determined easily from the transmission across a single layer,

$$\begin{matrix} V_n \\ I_n \end{matrix} = \begin{bmatrix} & \\ {}_n T_1 & \\ & \end{bmatrix} \begin{matrix} V_1 \\ I_1 \end{matrix} \tag{1.5}$$

$${}_n T_1 = {}_n T_m \; {}_m T_\ell \cdots {}_3 T_2 \; {}_2 T_1$$

$${}_j T_i = \text{Transmission across } i\text{-}j \text{ layer}$$

Thus the statistical parameters of ${}_n T_1$ are simply related to the statistical parameters of the network elements. To solve for the conduction across the whole network one must invert a submatrix of T however, and the statistical properties of the inverse matrix are no longer simply related to the statistical properties of the original matrix.

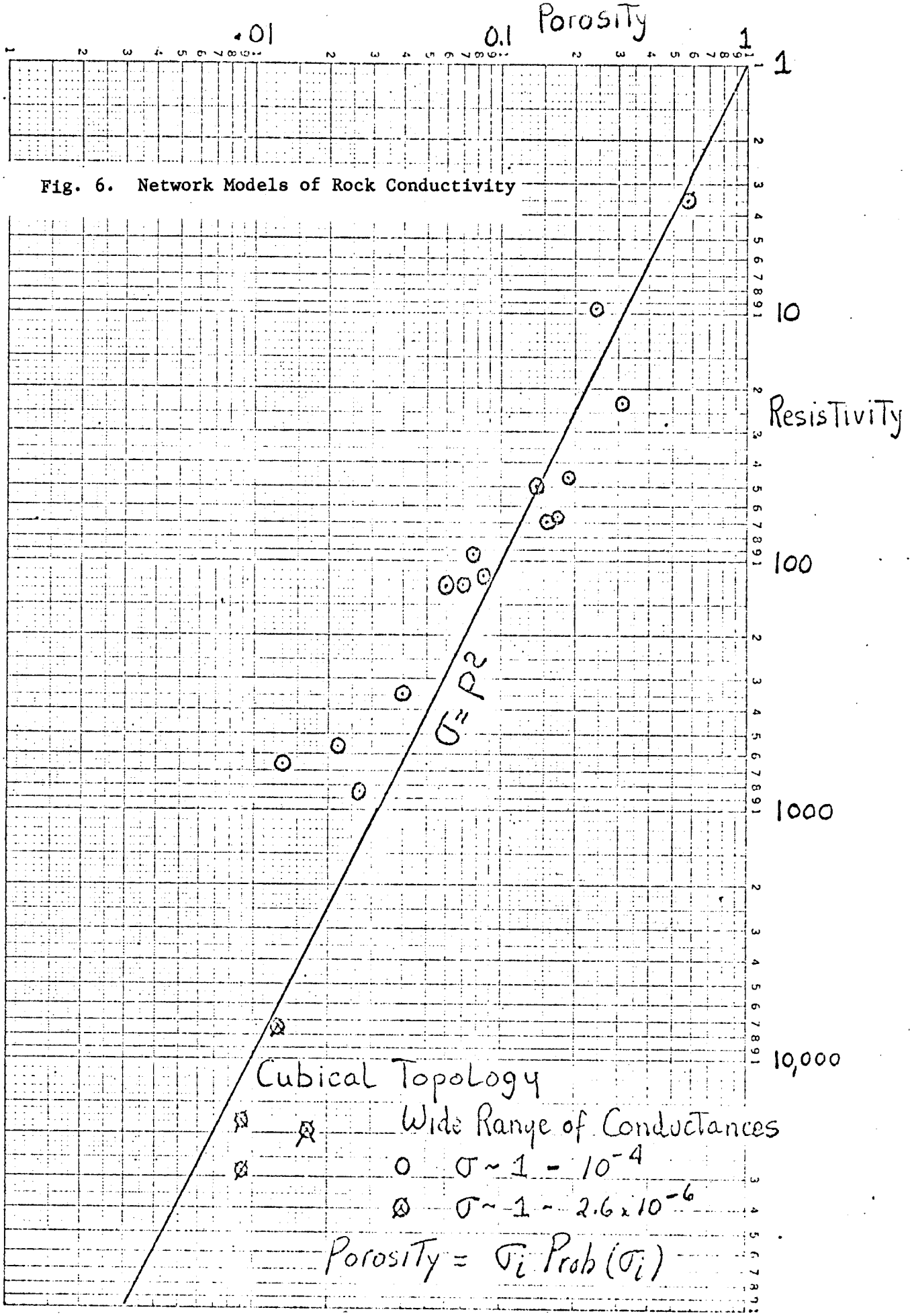


Fig. 6. Network Models of Rock Conductivity

Although this approach may be fruitful for one well versed in matrix theory it appeared too difficult to us and was abandoned.

At this point we tried a different approach which is illustrated in figure 7. If we choose a large enough rock segment it will appear to have homogeneous electrical properties and its conductivity can be represented by a single element of fixed value. The segment could be broken up into parts which are interconnected, but now the parts might show small statistical variations. This process could be cascaded until we had broken the rock down into elemental fragments which showed the full statistical variations of the crack and pore conduction properties. In order to compute the electrical properties of the rock one only has to work the process in reverse, starting with the elemental fragments with their full statistical variations and cascading up to the homogeneous average rock property. This sort of averaging procedure will be given the symbol << >> . If the second level is a large network of elemental segments its electrical properties will be close to the final answer. This is essentially what the computations whose results are shown in figure 6 amounted to. The complexity of a large network, however, makes analytic progress difficult. Therefore we use instead rather simple networks, but which preserve the features of series and parallel arrangements. Even very simple networks can have a fair number of elements, and in order to further simplify the problem, some elements can be made dependent elements. Figure 8 shows two very simple examples. The top example cannot be used to describe conduction from right to left, as the symmetry prevents any averaging being produced by the network. Because of the restrictions imposed on the elements we call these systems pseudo-random networks.

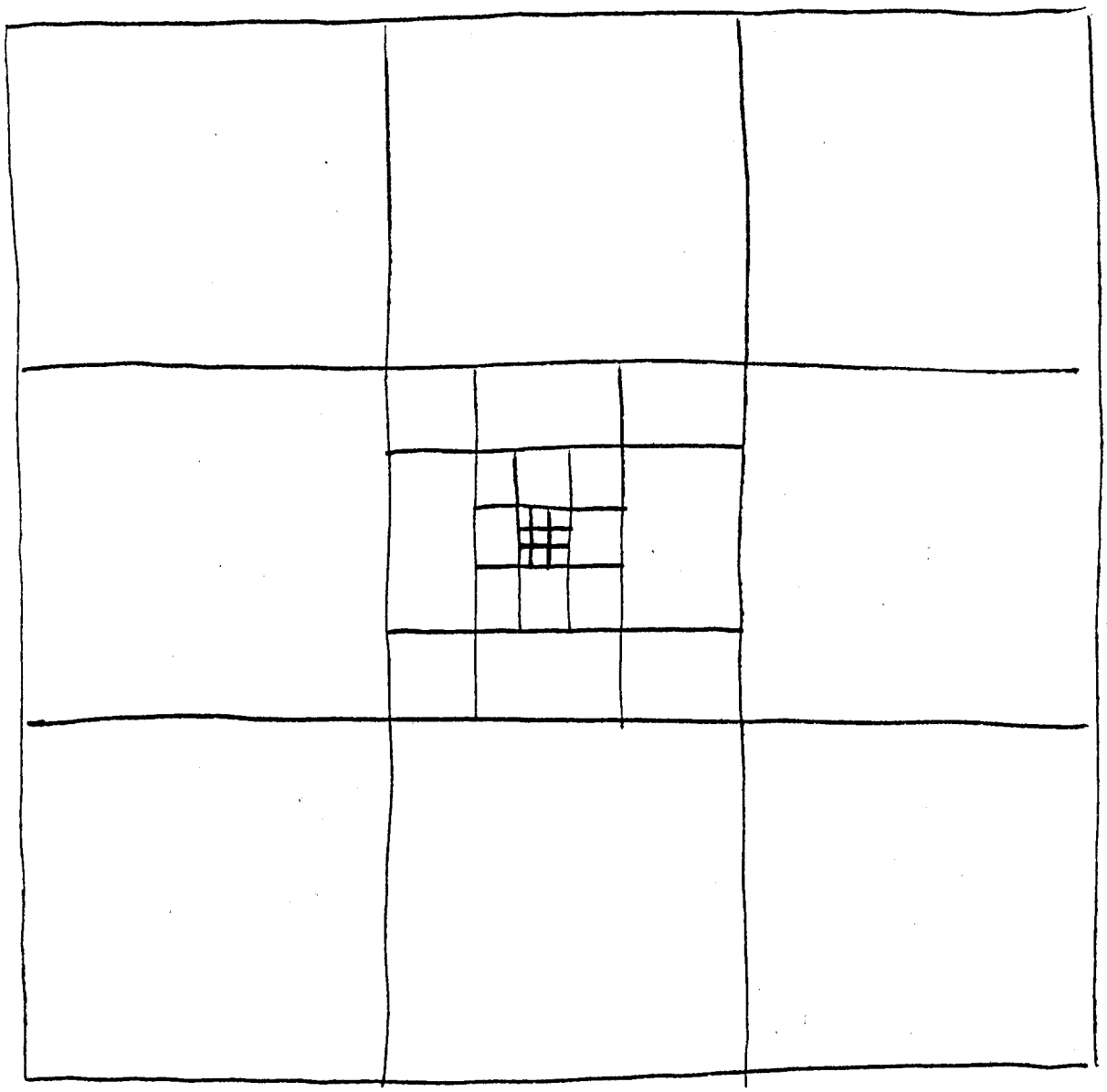
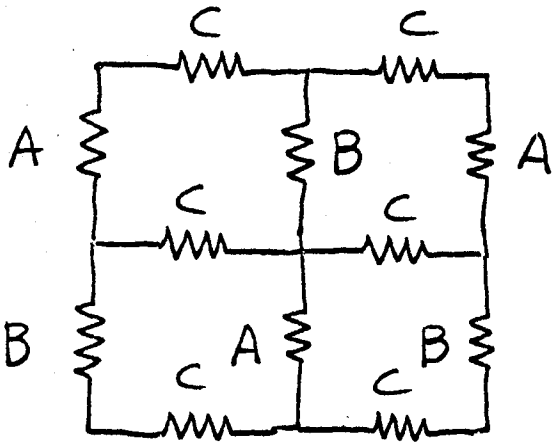
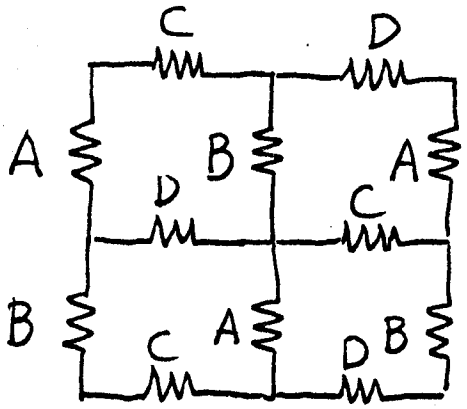


Fig. 7. Cascaded Network Subdivisions



Network I



Network II

Fig. 8. Examples of primary networks

Vertical Conductivity

$$\text{Network I} = \frac{2Y_AY_B + Y_C(Y_A+Y_B)}{Y_A + Y_B + 2Y_C}$$

$$\text{Network II} = \frac{2Y_AY_B + .5(Y_C+Y_D)(Y_A+Y_B)}{Y_A + Y_B + Y_C + Y_D}$$

The advantage of this sort of representation quickly became evident as certain quite difficult concepts became readily apparent. The concept of a critical probability for instance, is easily demonstrated. Let us consider a bimodal distribution for the conduction elements, their value being either 1 or 0. If the probability of an element having a zero conductance is Q_0 one can easily compute the probability of the first level network having zero conductance. For network I for instance the conductance is zero whenever

$$\begin{aligned} &Y_A, Y_B, Y_C = 0 && \text{prob} = Q_0^3 \\ &Y_A, Y_B = 0 \\ &Y_A, Y_C = 0 && \text{prob} = 3Q_0^2(1-Q_0) \\ &Y_B, Y_C = 0 \\ &\text{total prob } Q_1 = 3Q_0^2 - 2Q_0^3 \end{aligned} \tag{1.6}$$

The probability of the next level having zero conductance is therefore

$$\begin{aligned} Q_2 &= 3Q_1^2 - 2Q_1^3 \\ \text{and} \quad Q_3 &= 3Q_2^2 - 2Q_2^3 \end{aligned} \tag{1.7}$$

Since these functions are monotonically increasing between 0 and 1

$$\begin{aligned} &\text{if} \quad Q_n > Q_m \quad \text{then} \quad Q_\infty \to 1 \\ &\text{if} \quad Q_n < Q_m \quad \text{then} \quad Q_\infty \to 0 \end{aligned}$$

Thus a critical probability exists which is defined as

$$\begin{aligned} Q_c &= 3Q_c^2 - 2Q_c^3 && 0 < Q_c < 1 \\ &\text{or} \\ Q_c(Q_c-1)(Q_c-.5) &= 0 && 0 < Q_c < 1 \\ &\text{or} \quad Q_c = 0.5 \end{aligned} \tag{1.8}$$

This happens to be probably the correct critical probability for square networks. Network II gives a different result. It is open circuited for

$$
\begin{aligned}
&Y_A, Y_B, Y_C, Y_D && = 0 && \text{prob} = Q_0^4 \\
&Y_A, Y_B, Y_C \\
&Y_A, Y_B, Y_D && = 0 && \text{prob} = 4Q_0^3(1-Q_0) \\
&Y_B, Y_C, Y_D \\
&Y_A, Y_C, Y_D \\
&Y_A, Y_B && = 0 && \text{prob} = \frac{Q_0^2(1-Q_0^2)}{Q_1 = Q_0^2 + 2Q_0^3 - 2Q_0^4} \qquad (1.9) \\
& && && Q_c \approx .70
\end{aligned}
$$

More complicated networks lead to higher order rational fractions for the conductance formula, but they all have in common that they are homogeneous to the first degree in Y and their minimum and maximum values are equal to the minimum and maximum values of the Y's. Thus the higher level network distribution functions are confined to within the same limits as the original element values, but the probability of attaining these limiting values becomes vanishingly small. Figure 9 shows the progression of conductance distribution functions as we go to higher and higher levels.

The symmetry of the conductance equation for network I allows us to immediately obtain the final result in special cases and leads us to a useful approximation. If we denote $Y(Y_A,Y_B,Y_C)$ as the network conductance when the elements A, B, and C have conductances

PSEUDO RANDOM NETWORK INTERBEDDING

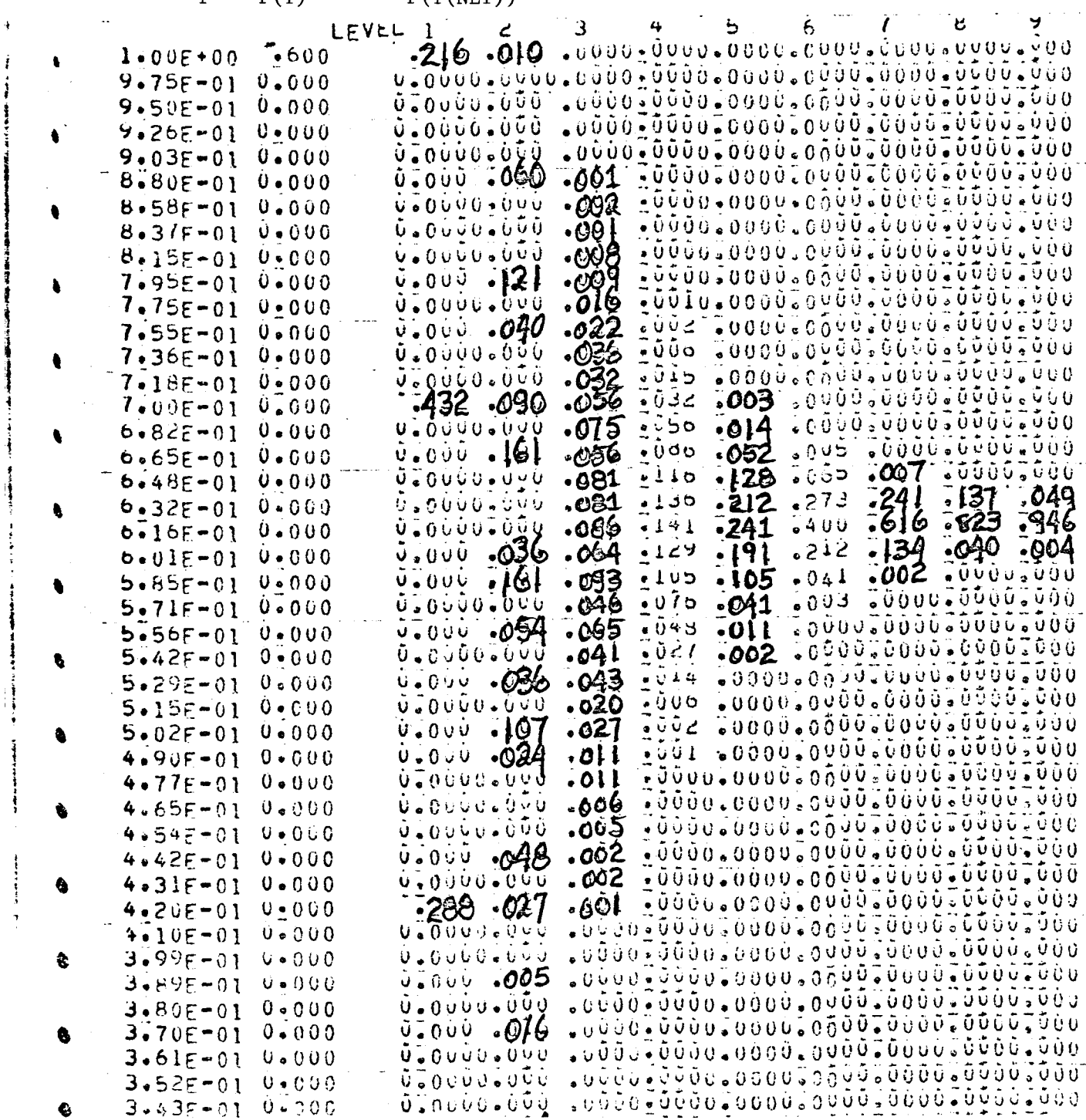


Fig. 9

Example of interbedded network distribution functions.

$$\prod y^{P_i} = .613$$

Y_A, Y_B, and Y_C then for network I

$$Y\left(\frac{Y_m}{A}, \frac{Y_m}{B}, \frac{Y_m}{C}\right) = \frac{Y_m^2}{Y\,(AY_m, BY_m, CY_m)} \tag{1.10}$$

If the distribution function of the element conductances is symmetric about Y_m on a logarithmic scale so that $P(AY_m) = P(Y_m/A)$ then the geometric mean of the conductances is Y_m,

where Geometric mean of $Y = \Pi\, Y_i^{Pi}$ (1.11)

since in this case (1.11) reduces to $Y_m^{\Sigma Pi}$

Because of (1.10) this same symmetry is retained in the network distribution function, and thus the geometric mean of the network is also equal to Y_m. This continues to hold for all the higher levels, and as the distribution functions tend towards a delta function at high levels Y_m must be the final network conductance. Even when the distribution function is not symmetric on a log scale, the geometric mean is a good approximation as long as the spread of values is not too great.

Let us condiser a bimodal distribution for the conductances, with values of 1 and Y and probabilities p and (1-p). The distribution function for the network (network I) is therefore

$$\begin{aligned}
P\,(1) &= p^3 + p^2\,(1-p) = p^2 \\
P\left(\frac{Y(1+3Y)}{Y+3}\right) &= 2p^2\,(1-p) \\
P\left(\frac{(y+3)}{1+3Y}\right) &= 2p\,(1-p)^2 \\
P\,(Y) &= (1-p)^3 + p(1-p)^2 = (1-p)^2
\end{aligned} \tag{1.12}$$

The geometric mean of the element conductance values is Y^{1-p} while the geometric mean of the network conductance values is

$$Y^{(2p+1)(1-p)^2} \left(\frac{3 + Y}{1 + 3Y}\right)^{2p(1-2p)(1-p)} \tag{1.13}$$

These values are identical at $p = 0, .5,$ and 1 but they track very closely as long as Y does not deviate too far from 1. If we consider $p = .25$, which should represent close to the worse case, we find very little error up to $1{:}Y$ values of three. Table II shows these results

Table II

Comparison of geometric mean of conductances and network I for element values 1, Y and $P(1) = .25$

1:Y	1.1	1.3	2	3	5	10
% difference of geometric means	0	0	.4	.9	1.8	5.7

If the same analysis is done for network II it is found that the network geometric mean and the element geometric means are equal at $p = 0, p_c, 1$. It can probably be shown that this result holds for all networks. Network II deviates more from the element mean than network I in the region $p>p_c$ and in this region its mean is higher than the element mean.

Even when the spread of values is too great to be able to use the geometric mean as a final answer, one can use the concept to greatly reduce the computations involved in cascading up to higher levels. The computational load arises from the fact that the number of different values that the network can have increases geometrically with the level. This effect can be circumvented by, at each step,

combining groups of values within some range into a single value with a combined probability. If the geometric mean of the group is used, very large ranges can be combined together with very little loss of accuracy, but a tremendous saving in computational time.

Even though the geometric mean loses its accuracy when the spread of values becomes large it is still a very useful approximation since usually in such cases high accuracy is not necessary. Thus if we consider a distribution function with probability .1111 for Y values of

1, .32, .1, .032, .01, .0032, .001, .00032, .0001

the geometric mean is .01 and network II cascades to a value of .018. These differences do not seem too dramatic when plotted on 4 cycle log paper.

Since networks can be used as analogues of difference equations for many field equations, these same ideas can be applied to finding the physical properties of heterogeneous media. Figure 10 shows data on the elastic properties of a We-Co mix. Also shown are the results of random networks and cascaded pseudo-random networks. The inaccuracies of the experimental data are too great to provide a critical test, but it would appear we are on the right track.

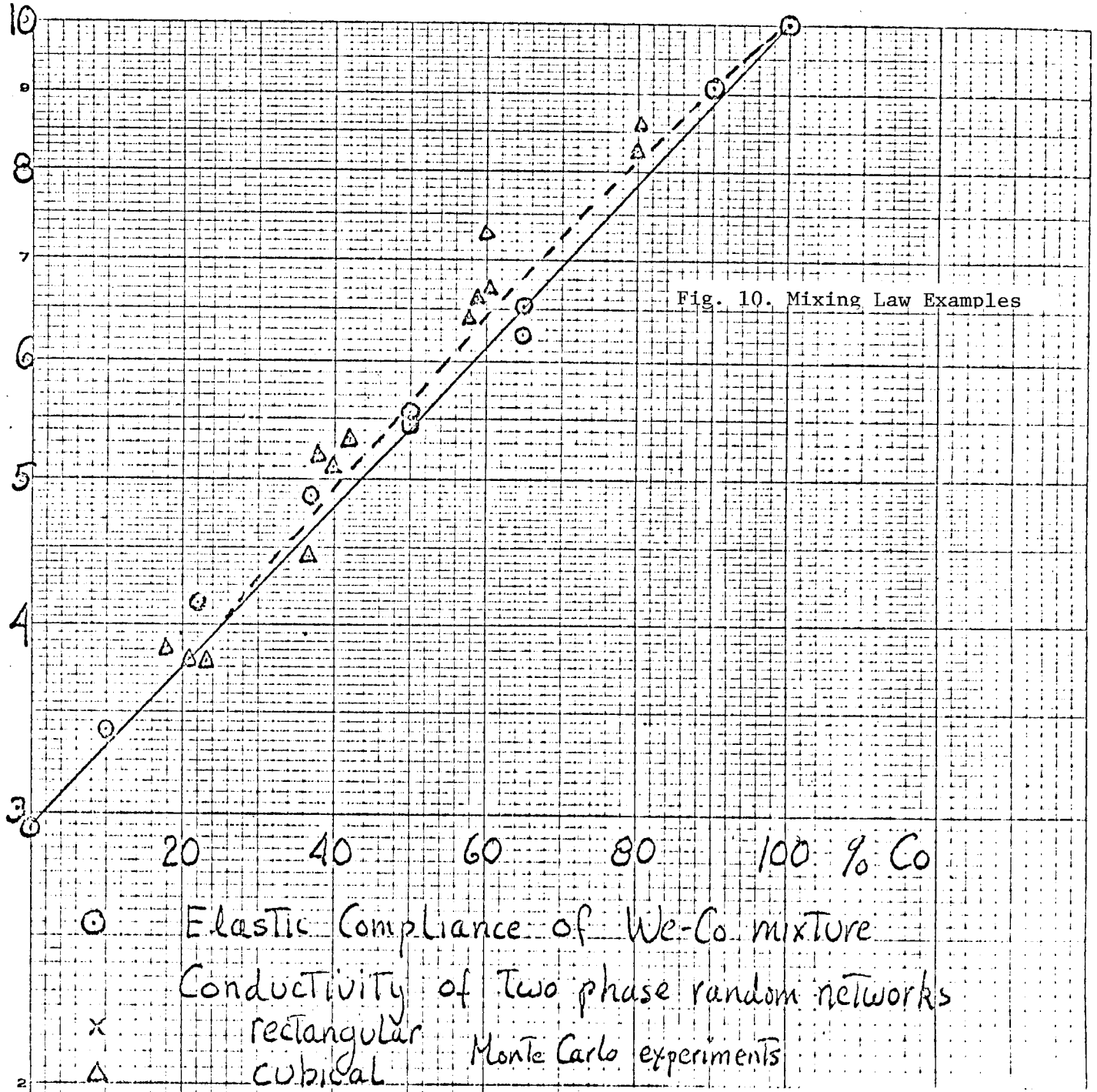


Fig. 10. Mixing Law Examples

The rectangular networks and the psuedo-random net $Y = \ll \frac{2Y_AY_B + Y_C(Y_A+Y_B)}{Y_A+Y_B+2Y_C} \gg$ followed the geometric mean

Let us return again to the question of Archie's law and its implications using the ideas we have been outlining. It is clear that in general one cannot expect a law like Archie's law to hold unless some restrictions are placed on the distribution function, except perhaps at the limit of very high porosities. If we consider network I and take a bimodal distribution of element values

$$\begin{aligned} Y = 1 \quad & P(1) = p = (1-q) \\ Y = 0 \quad & P(0) = 1-p = q \end{aligned} \tag{1.14}$$

then network I has the following distribution of conductances

$$\begin{aligned} P(Y = 1) &= (1-q)^2 \\ P(Y = 1/3) &= 2q(1-q)^2 \\ P(Y = 0) &= q^2(3-2q) \end{aligned} \tag{1.15}$$

From these we compute $\langle Y \rangle = 1 - 1.33q - .33q^2 + .67q^3$ (1.16)

At the next level we have 9 different values for Y, but keeping only terms to second order in q we have

$$\langle Y \rangle \text{ second level} = 1 - 1.6q - .63q^2$$

In table II we show how these results seem to indicate a trend towards $\langle\langle Y \rangle\rangle = (\text{porosity})^2$.

Table II

High porosity expansions of Pseudo Random Network Sequence

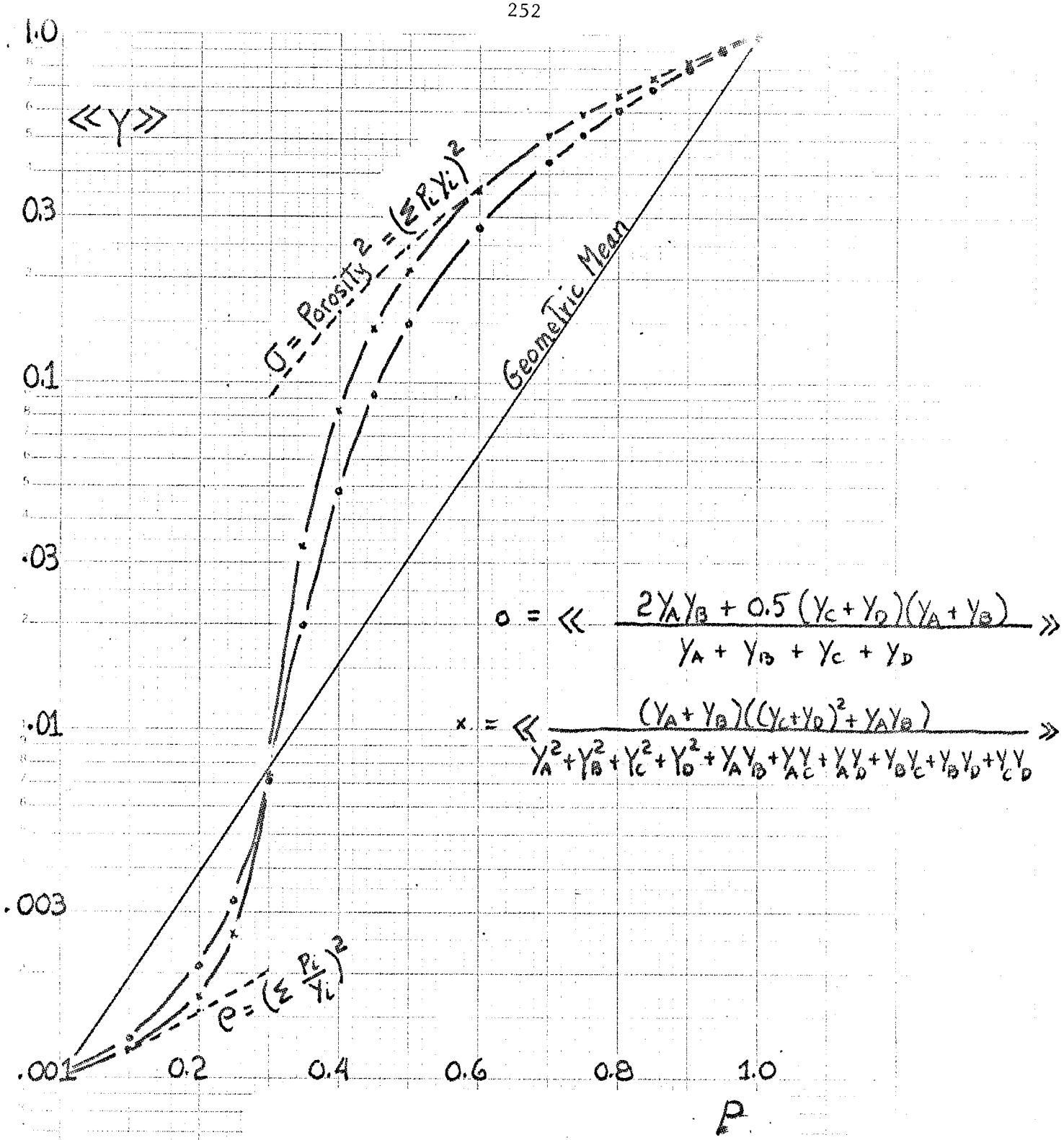
for Network I

Prob of open circuit	Porosity	$(\text{Porosity})^2$	$< Y >_2$	$< Y >_1$
.01	.99	.98	.984	.987
.03	.97	.941	.951	.960
.1	.90	.81	.834	.864
.3	.70	.49	.463	.571

This trend does not appear to be a function of the network topology if we avoid extreme topologies. As we mentioned before each topology has a critical probability P_c, and this parameter appears to closely limit the behavior of the network when random element values are used. In figure 11 we show the behavior of two different pseudo random networks with equal critical probabilities. One network is network II which was shown in figure 8. The other network has not been synthesized. It is more complicated than network II, although it has the same number of independent element values and the same critical probability. Both networks follow the porosity squared law at high porosity, but the more complicated network maintains this trend to lower probabilities. Both networks merge with the geometric mean at the critical probability and below that appear to fit a trend which is a sort of dual of Archie's law

$$\frac{1}{\sigma} = \left(\Sigma \frac{P_i}{Y_i} \right)^2 \quad \text{for low porosity} \tag{1.17}$$

It is very tempting to infer that very large networks would follow the two trends right up to the critical probability where they would jump from one



$P(y=1) = p$

$P(y=.001) = 1-p$

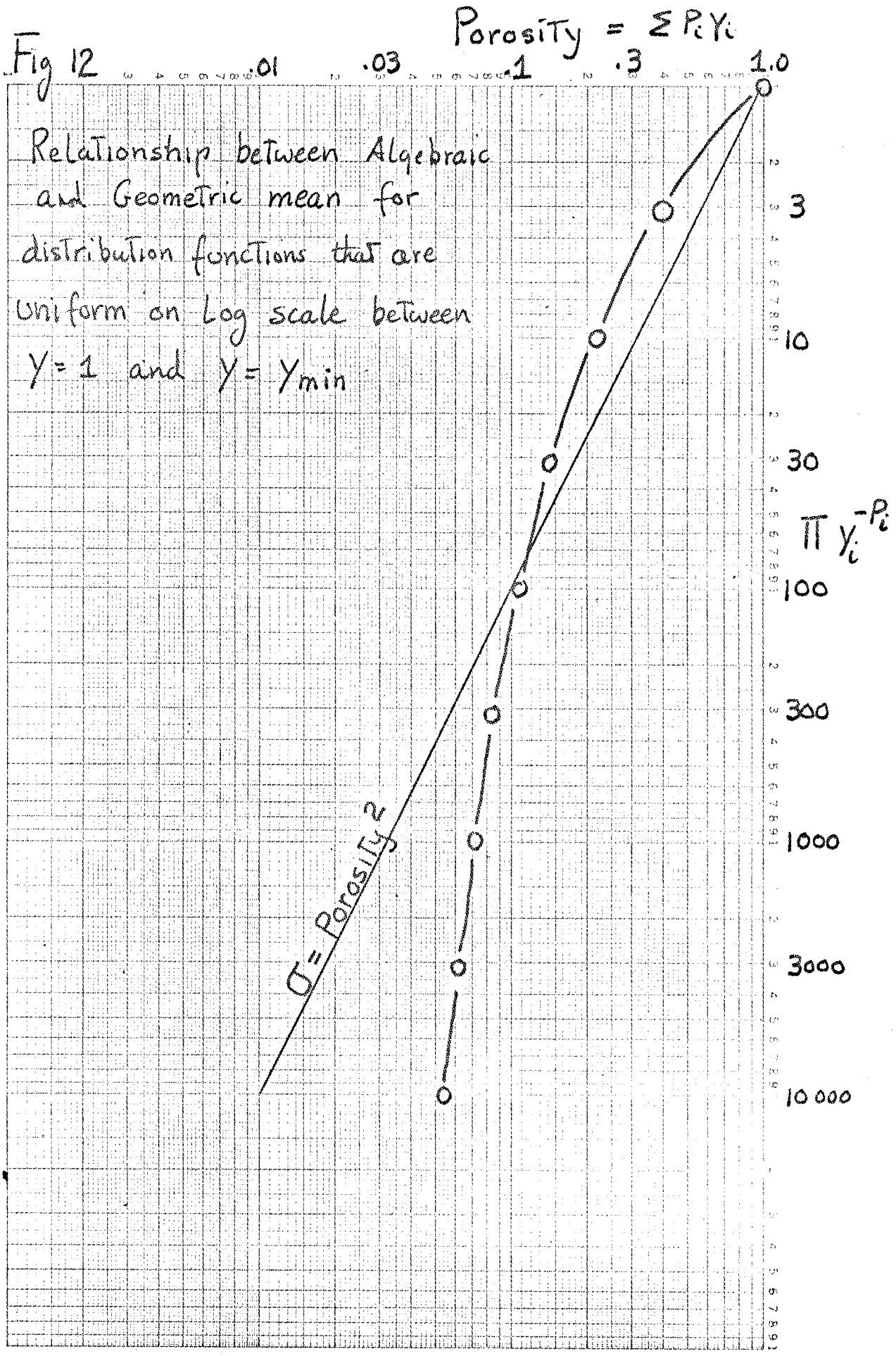
Fig 11 Comparison of Pseudo Random networks with similar critical probabilities

trend to the other one. As yet we have not examined more complicated pseudo random networks in order to verify this inference.

The differences between the two networks shown in figure 11 are greatly reduced when smoother distribution functions are used. With a probability distribution function of $P_i = .1111$ for $Y_i = 1.0, .32, .1, .032, .01, .0032, .001, .00032, .0001$ the two networks gave final answers differing by only 4%.

The $(\text{porosity})^2$ behavior that we have been examing thus appears to be independent of the network topology, but it is not a universal behavior since it depends on the form of the distribution function. In the cases just discussed we were using a bimodal distribution of Y values, which two values were widely separated. If instead we use a distribution function which has only a very short spread of Y values we find a porosity to the first power law. This can be proven easily since in these cases the geometric mean is an accurate approximation. If we consider uniform distribution functions on a logarithmic scale and use the geometric mean as an approximate answer we obtain the results shown in figure 12. At modest porosities the results, though not a $(\text{porosity})^2$ law, do not deviate too far from Archie's law, but as the minimum Y_i is extended to smaller values a drastic decrease in the geometric mean occurs relative to the porosity. At this end of the plot the geometric mean cannot be expected to remain an accurate approximate, but nevertheless one realizes that Archie's law behavior at low porosities must involve limitations on the conductivity parameter distribution function. Actual distribution functions must lie in between the two extremes of a very narrow distribution function and one uniformly distributed over a wide range (on a log scale). A distribution function which involves a linear scale is bound to be too narrow to explain the observations for low porosity rocks. Distribution functions like the

Fig 12



Relationship between Algebraic and Geometric mean for distribution functions that are uniform on Log scale between $Y = 1$ and $Y = Y_{min}$

the log normal should work provided some relationship between the mean and the deviation is maintained. We have not worked out such relationships, nor have we looked at the data on crack and pore width distributions.

Several extensions of this work are needed in order to more completely understand these mixing laws. One simple extension is to consider anistropic distribution functions. A more difficult and subtle extension is one needed to understand the role of shape factors. This may not be as important in the resistivity problem as it is in other mixing problems. Since networks can represent difference equation approximations of field equations, in the limit of small enough grid spacings they should give correct answers. How then can we reconcile the fact that different network topologies give different answers. The reason for this is that when we assigned the distribution function to the network elements we assumed each element was independent. This is roughly equivalent to thinking that each element represented a single grain of the rock matrix, and thus the network topology used was actually implying something about the grain geometry. To become independent of the network topology one must use a gird spacing smaller than the grain size, but then the individual elements are no longer independent. This greatly complicates the calculations, but it allows one to deal with such important factors as the actual distribution of shapes and also eliminates the bias introduced by a choice of network topology. The use of very simple pseudo random networks still involves an approximation, but the errors introduced by this will be reduced much as these errors were reduced when dealing with very broad distribution functions.

Curve Fitting in Geophysical Inverse Problems

by Jon F. Claerbout

(Mini-lecture given September 12, 1973 based on research done in Australia)

Let's call the following functions "model norms."

$$N_1(y, \bar{y}) = |y - \bar{y}| \qquad (1\,a,b,c,d)$$

$$N_2(y, \bar{y}) = (y - \bar{y})^2$$

$$N_3(y, \bar{y}) = -\bar{y}\,\ell n\,(y/\bar{y}) + y - \bar{y}$$

$$N_4(y, \bar{y}) = y\,\ell n\,(y/\bar{y}) - y + \bar{y}$$

These model norms share the property that they are minimized at $y = \bar{y}$.

Let the continuum of earth properties be mapped into a vector x and the finite number of data measurements be mapped into a shorter vector d. Occasionally we have linear problems (more often we linearize non-linear problems). These are indicated by the constraint equations

$$\begin{bmatrix} A \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} d \end{bmatrix} \qquad (2)$$

To relate (2) to (1) we have a function f.

$$y = f(x_1, x_2, \ldots, x_m) \qquad (3)$$

Often the function f is taken to be simply $y_i = x_i$. To do inverse problems (curve fit) we generally find the model x which minimizes

$$\min_{x_i} = \sum_i w_i N(y_i, \bar{y}_i) \qquad (4)$$

In the absense of the constraint equations (2), the minimization of (4) should lead to $y = \bar{y}$ which should imply $x = \bar{x}$ where $\bar{x}$ is called the "initial model" or the "default model". Note that it is not necessary to specify the default model $\bar{x}$ directly, but it is necessary to specify $\bar{y}$. For example if f says that y is the spatial gradient of x then choice of $\bar{y} = o$ specifies that a homogeneous earth is the default model, but it does not specify the space independent numerical value of $\bar{x}$.

Obviously there is a great deal of arbitrariness in the choice of the model norm N, the choice of the function f, the choice of the default $\bar{y}$, and the choice of a weighting function w. Ordinarily this arbitrariness must be resolved by "geophysical intuition". However, there are a few general principles which are helpful.

My purpose today is (1) to show you that the use of N_3 is like maximum entropy spectral estimation; (2) to indicate that N_4 will have advantages over N_3 in some applications; and (3) to advocate certain restrictions on the choice of N, f, and w which I believe will then lead to smooth high resolution models. People familiar with maximum entropy spectral estimation are aware of its astounding ability to resolve tiny spectral peaks in the vicinity of massive spectral peaks. There is no "Gibbs phenomena". The present study is motivated by the desire to bring this combination of stability and resolving power to other applications.

To begin with, I will skim over a proof that Toeplitz equations (the most common case of maximum entropy) result from N_3 with $w = 1$ and $\bar{y} = 1$. Given an input power spectrum R we seek a filter with power response S so that the output spectrum $y = RS$ should tend to the constant $\bar{y} = +1$ which is independent of frequency. When this is achieved the input spectrum R is deemed to be $1 / S$. The minimization (4) with N_3, $w = 1$ and $\bar{y} = 1$ becomes

$$\min = \int - \ell n\, RS + RS \; d\omega \qquad (5)$$

Now S is a function of the parameters in the x vector which we will subsequently define. Thus, to achieve the minimum in (5) we will require zero partial derivatives of (5) with respect to each x_i. Thus for each i we have

$$0 = \int \left(-\frac{1}{S} + R\right) \frac{\partial S}{\partial x_i} \; d\omega \qquad (6)$$

Let us define a minimum phase filter

$$A(z) = a_0 + a_1 z + a_2 z^2 + \cdots$$

3.

its inverse,

$$B(z) = 1/A(z) = b_0 + b_1 z + b_2 z^2 + \cdots$$

the relation between z-transforms and fourier transforms $z = e^{i\omega}$, and the spectrum

$$S = \bar{A}(e^{-i\omega}) \; A(e^{i\omega})$$

The adjustable parameters a_i are sufficient to make any continuous spectrum S. Therefore, we can take the spectrum to be parameterized by $x_i = \bar{a}_i$. The problem becomes more manageable if we introduce the constraints $\bar{a}_i = o$ for $i > N$. Now we have for the required partial derivative

$$\frac{\partial S}{\partial x_k} = \frac{\partial S}{\partial \bar{a}_k} = z^{-k} A(z) \qquad (7)$$

Inserting into (6) we get

$$o = \int e^{-ik\omega} \left(- \frac{1}{\bar{A}(e^{-i\omega k})} + R'A \right) d\omega \qquad \text{for } k = o,n$$

or

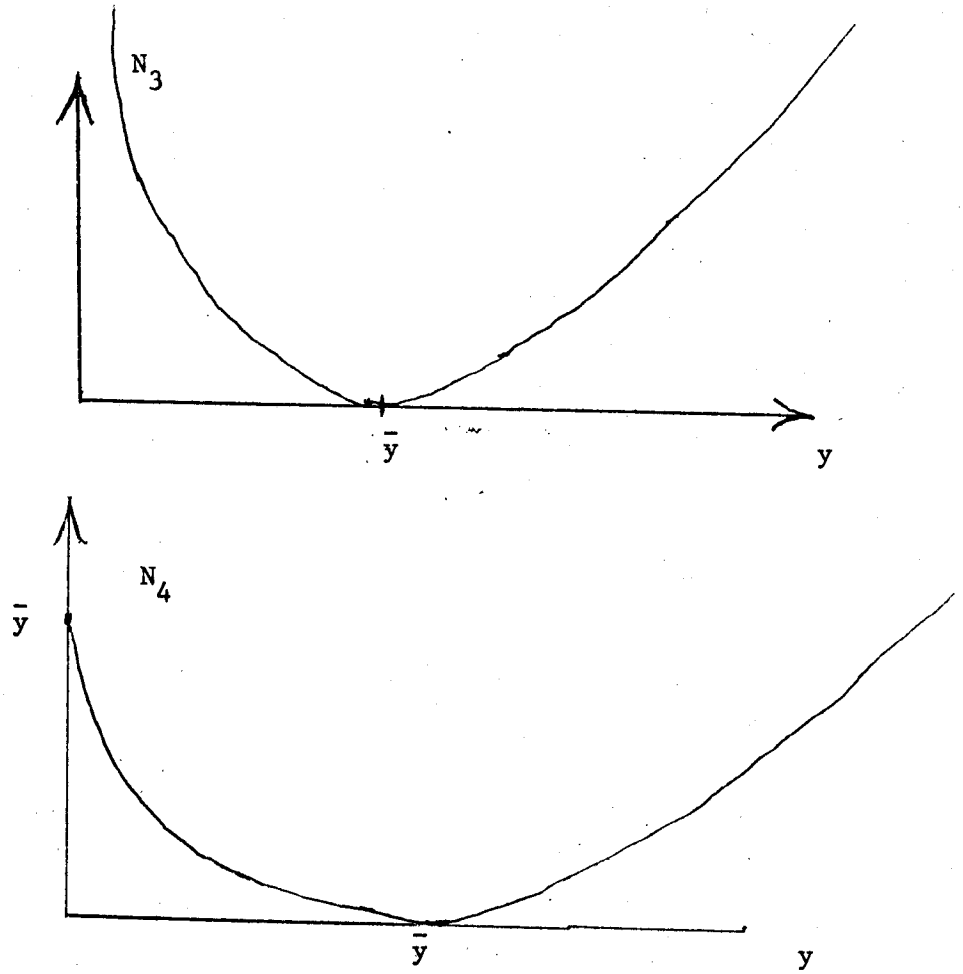
$$o = \int z^{-k} \left(-(\bar{b}_0 + \bar{b}_1/z + \cdots) + R\,A \right) d\omega$$

Since this integral selects the coefficient of z^o of the argument we obtain for $n = 3$

$$\begin{bmatrix} r_o & r_{-1} & r_{-2} \\ r_1 & r_o & r_{-1} \\ r_2 & r_1 & r_o \end{bmatrix} \begin{bmatrix} a_o \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} b_o \\ o \\ o \end{bmatrix} \qquad (8)$$

4.

The equations (8) may be deduced in numerous ways. One of John P. Burg's contributions was to show, similar to the way I just indicated, that (8) is a consequence of the use of N_3. Now I wish to plot N_3 and N_4. They look like



Their derivatives are

$$\frac{\partial N_3}{\partial y} = -\frac{\bar{y}}{y} + 1 \tag{9}$$

$$\frac{\partial N_4}{\partial y} = \ell n\ y/\bar{y} \tag{10}$$

and their second derivatives are

$$\frac{\partial^2 N_3}{\partial y^2} = \bar{y}/y^2 \tag{11}$$

$$\frac{\partial^2 N_4}{\partial y^2} = 1/y \tag{12}$$

Clearly the first derivatives vanish as they should at $y = \bar{y}$ and the second derivatives are positive (as they should be for convexity) for all positive y. Careful inspection reveals that N_4 is finite at $y = 0$ but has infinite slope there.

Now we can think about how seismic deconvolution is always done with N_3, with $w = 1$ and $\bar{y} = 1$ and enquire whether some of the other possibilities might not be preferable. First of all $\bar{y} = 1$ independent of frequency is bad because it whitens the output forcing us to bandpass the data after decon. Although we might like to use other $\bar{y}$ functions for the desired output spectrum it turns out that they destroy (8) and lead to a mess.

It doesn't mean we couldn't do it; it just means it would cost more. Second, N_4 as compared with N_3 will concern itself more with knocking down peaks and less with drawing up the holes. In the limit of very many filter points both N_3 and N_4 will give a white spectral output. The differences occur when the filters are constrained to be short. I think that deconvolution works because the spectral peaks get knocked down. Drawing up the holes is actually a disadvantage if they contain only noise.

Another fact about N_4 is that its derivative (10), which represents the "force" a given model point applies against other model points and against the constraints is symmetric about $\bar{y}$ in the following sense: if y were ten times as big as $\bar{y}$ then the force is equal and opposite to what it would be if y were ten times as small as $\bar{y}$. In the simple harmonic oscillator, force is linearly proportional to the distance from equilibrium. With N_4 the force is linearly proportional to excursion of the logarithm from equilibrium.

A central problem in geophysical curve fitting is to determine whether small bumps and irregularities on the curves really represent the earth or whether they just result from the arbitrariness in the choice of N, f, and w. A common way to suppress the bumps and irregularities is to choose f to be a spatial gradient or some high pass filtering function of the model. The trouble with such an approach is that it suppresses not only the bumps which result from a poor choice of N, f, and w but also the bumps which the earth really has. What we are after is a scheme to suppress the artificial bumps without suppressing the real ones. The following scheme is proposed:

First you must identify a "density" in the problem at hand. Densities have the property that their volume integral means something. For example, in a resistivity problem a density is the power dissipated per unit volume in the earth. In other problems we have mass density, kinetic energy density, potential energy density, probability density, and power spectral density. Then choose f so that $y = f(x)$ is the density function you have identified. (the resistances are the x_i .) Next the model norm function must be chosen to be a function of homogeneous type. Such functions commonly occur in thermodynamics and are defined by

$$1/2 \; N(y, \bar{y}) = N(y/2, \bar{y}/2) \tag{13}$$

(In a more precise definition the 1/2 is replaced by an arbitrary constant). Notice that this includes N_1, N_3, and N_4 but excludes N_2. Furthermore, all the weight factors w_i should be equal, say equal +1.

Because the summation (4) approximates a volume integral the extra points which arise from over-zealous sampling of y_i in one region of space have compensatingly small volumes associated with them. Our results should be independent of whether the unknowns x_i (which represent conductivity, compressibility, mass density, etc.) are stated on a cartesian grid, a cylindrical grid, or as coefficients of an expansion in any complete set of functions.