Title
THREE-DIMENSIONAL DENSITY RECONSTRUCTION FROM A SERIES OF TWO-DIMENSIONAL PROJECTIONS

Author
Goitein, Michael.

Publication Date
1971-12-01

Submitted to Nuclear Instruments
and Methods

LBL-547
Preprint
C.1

THREE-DIMENSIONAL DENSITY RECONSTRUCTION FROM A SERIES OF TWO-DIMENSIONAL PROJECTIONS

Michael Goitein

December 1971

AEC Contract No. W-7405-eng-48

DISCLAIMER

THREE-DIMENSIONAL DENSITY RECONSTRUCTION FROM A SERIES OF TWO-DIMENSIONAL PROJECTIONS*

Michael Goitein

Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720, USA

December 1971

ABSTRACT

An iterative relaxation technique is presented for unfolding three-dimensional distributions from a series of two-dimensional projections taken at several different orientations relative to the object being investigated. The achievable resolution is discussed. Analyses of both computer simultations and actual measurements are presented. This type of analysis has applications in radiography, electron transmission microscopy, radiotherapy, and nuclear medicine. Features of the technique which may make it especially suitable for particular applications are described.

* This work was done under the auspices of the U. S. Atomic Energy Commission.

1. INTRODUCTION

A series of x-ray pictures taken at a number of different view angles may contain enough information to enable one to reconstruct the full three-dimensional distribution of absorption coefficients in the viewed object. Similarly, a series of scans involving translations and rotations of a pair of gamma detectors may be used to map out the distribution of a positron-emitting isotope through a transverse section of a patient. On a quite different scale, transmission electron micrographs taken at a series of tilt angles may be used to reconstruct the three-dimensional underlying structure. Finally, measurement of the average energy loss of a high-energy heavy-particle beam passing through a patient leads to knowledge of the projected stopping power along the beam line. A series of such measurements at different transverse positions and patient orientations can be used to reconstruct the distribution of stopping power throughout a transverse section, and this information can be directly used to guide Bragg-peak radiotherapy.

These diverse situations present the same computational problem. We offer here a new method of analyzing the problem, using an iterative relaxation technique. A discussion of the relative merits of this and the other available mathematical models[1-3]) is deferred until after the description of the method.

It is self-evident that the three-dimensional problem can be broken up into a series of two-dimensional problems by considering separately a series of "slices" through the scanned object. We make this simplification in all the computational examples presented here and in our treatment of resolution. However, nothing in the mathematics of the analysis

implies such a limitation.

The basic series of measurements is illustrated schematically in fig. 1. It consists of a series of scans taken at a number of different angles relative to the object to be examined. Each scan comprises a series of measurements of the projected density along a number of discrete transversely separated lines. Of course, in some applications, the scan is almost continuous. In such a case we assume the information is "binned" into a number of discrete measurements.

In the technique to be described there is nothing which requires the scan to be taken at regular intervals of either angle or position. Indeed, there is no requirement that the scan-lines be parallel with one another, or even that the measured projections be along straight lines. However, for descriptive convenience, we will discuss the problem as though these conditions hold.

First, we must establish some notation. We divide the space upon which measurements are to be made into a region within which are N cells of unknown density, and outside of that region the density is assumed to be known exactly. That being the case, the contribution from the known density region can always be calculated and subtracted from the measurement, a situation equivalent to having a density of zero outside the region of unknown density. This is assumed to be the case in all that follows.

Within each cell the density is assumed to be uniform. This assumption could be modified, as is discussed later. The density of the ith cell is denoted ρ_i. We consider the N cells as being partitioned in a Cartesian grid with $n_x \times n_y$ divisions as depicted in fig. 2. This

particular grid is chosen only for simplicity. Other, more complicated cell patterns can easily be accommodated.

By a "measurement" we mean the result of measuring the line integral (projection) of density along a single path. A total of M such measurements comprises a complete scan. In most examples considered here there are m_x transversely displaced measurements at each of m_a orientations, and then $M = m_a \times m_x$. The value of each measurement is x_j $(j = 1, M)$ and the associated measurement error (standard deviation) is σ_j. We denote the theoretical value that the measurement should have (on the basis of some assumed density distribution) as X_j (x_j is the result of measurement with attendant errors). It is assumed to be related to the densities through the linear relationship

$$X_j = \sum_{k=1}^{N} f_{jk} \rho_k \qquad (j = 1, M). \tag{1}$$

This linear relationship may not, of course, hold for the primary measurements made. In that case the X_j are to be interpreted as secondary quantities derived from the measurements. For example, in a counting detector used in gamma-ray transmission measurements the X_j would represent the logarithm of the count rate (normalized to the rate without absorber).

A typical measurement swathe is depicted schematically in fig. 2. The interpretation of the coefficients f_{jk} of eq. (1) is that they are the average path length within the kth cell of the jth measurement. The average is taken, properly weighted, over the beam profile so as to take into account the effect of beam width. Many, indeed most, of the f_{jk} will be zero for reasonably narrow beams.

2. EXACT SOLUTION

We thus have a linear problem in N unknowns (the N ρ_i). The number of degrees of freedom would appear at first sight to be M - N, but there is a correction due to the correlation between sets of measurements at different angles. The sum of all measurements at a given angle is a constant independent of that angle. Thus there are m_a-1 constraint equations and, hence, $D = M - (m_a - 1) - N$ degrees of freedom. Provided $D \geq 0$ there can be a solution. The least-squares solution is that for which

$$\mathcal{M}^2 = \sum_{i=1}^{M} \frac{(X_i - x_i)^2}{\sigma_i^2} \tag{2}$$

is a minimum. It is trivial to show that this condition is met by the solution of the N simultaneous linear equations in N unknowns:

$$\sum_{k=1}^{N} a_{ki} \rho_k = b_i \quad (i = 1, N),$$

where

$$a_{ki} = \sum_{j=1}^{M} \frac{1}{\sigma_j^2} f_{jk} f_{ji}$$

and

$$b_i = \sum_{j=1}^{M} \frac{1}{\sigma_j^2} X_j f_{ji} .$$

These equations can be solved by standard matrix inversion techniques. We term the resulting solution the "exact solution," meaning that it is the result of a unique specification with a well-defined unique solution. The trouble is, of course, that the solution of these equations involves the inversion of an N×N matrix. For most interesting subdivisions this imposes impossible demands on both core and

time availability in even the largest computers. For example, dividing the object into a modest 15×15 grid would entail inversion of a 225×225 matrix. This array alone would require more than 50 000 words of core for storage. Moreover, even though techniques are available for handling matrices which overflow core, it must be remembered that execution time goes up with roughly the third power of the matrix, hence, with the sixth power of the number of cells along the edge of the object.

3. ITERATIVE RELAXATION TECHNIQUE

The solution we have developed to meet the computational inaccessability of an "exact solution" involves an iterative procedure. At the start of any given iteration one has a density value assigned to each cell. Consider the ith cell and all those projections which involve a contribution from it, that is, the small fraction of all the measurements for which $f_{ji} \neq 0$ $(j = 1, M)$. The heart of the technique is to adjust the density of the ith cell in such a way as best to fit all measurements which involve that cell. All other cells are assumed to have the fixed values assigned at the start of the iteration. The "best fit" is judged on the basis of a least-squares minimization. Specifically, we may rewrite eq. (1):

$$f_{ji}\rho_i = X_j - \sum_{\substack{k=1 \\ k\neq i}}^{N} f_{jk}\rho_k \qquad (j = 1, M).$$

This gives M equations in one unknown (ρ_i), although, since many f_{ji} will be zero (many measurements have nothing to say about a specific cell), there will be far less than M interesting equations.

We find the solution in the least-squares sense by requiring that

$$\mathcal{m}^2 = \sum_{l=1}^{M} \frac{(X_l - x_l)^2}{\sigma_l^2} \quad \text{be a minimum.}$$

This will occur for $d\mathcal{m}^2/d\rho_i = 0$. The solution, which involves only trivial algebra is then:

$$\Delta\rho_i = \rho_i^{n+1} - \rho_i^n = \frac{\sum_{l=1}^{M}\left(\frac{f_{li}}{\sigma_l}\right)\frac{1}{\sigma_l}\left(x_l - \sum_{k=1}^{N} f_{lk}\rho_k^n\right)}{\sum_{l=1}^{M}\frac{f_{li}^2}{\sigma_l^2}} \tag{3}$$

where ρ_i^{n+1} is the adjusted cell density and ρ_i^n is the density assumed at the start of the nth iteration.

3.1 Damping

One might assume that one could calculate the N adjustments ($\Delta\rho_i$; i = 1, N) and apply them to all cells without further manipulation. This procedure, however, is seriously deficient in that it leads to a rapidly diverging solution which blows up after only a few iterations. The reason for this behavior is easy to see. Consider the situation in which, on average, the cells have too low a density at the start of an iteration. As each cell is examined there will be a tendency to increase its density, over and above the particular adjustment required to improve the local density variations. This increase is made assuming all other cells have the value assigned at the start of the iteration and does not take into account the fact that they too will be increased to account for the overall low density. Thus when all cells are adjusted there will be a tendency to overcompensate for the overall density deficit. This

problem will clearly lead to increasingly large overshoots with successive iterations.

The solution which we have adopted to meet this problem is to introduce an overall multiplicative damping factor, δ. We then compute the densities used as input to the (n+1)th iteration (ρ_i^{n+1}) according to the formula

$$\rho_i^{n+1} = \rho_i^n + \delta \, \Delta \rho_i , \tag{4}$$

where $\Delta \rho_i$ is the quantity expressed in eq. (3).

One might envision many ways of achieving this damping effect. We have chosen δ by the requirement that the overall solution (involving all cells) be such that the measurements are best matched by the new densities in the least-squares sense. This is the requirement that

$$\mathcal{M}^2 = \sum_{j=1}^{M} \frac{(X_j - x_j)^2}{\sigma_j^2} \quad \text{is a minimum}$$

which occurs for $d\mathcal{M}^2/d\delta = 0$. The X_j are, of course, functions of δ through eqs. (1), (3), and (4). They may be expressed as

$$X_j = \sum_{k=1}^{M} f_{jk} \, (\rho_k^n + \delta \, \Delta \rho_k).$$

The solution involves a little algebra and may be conveniently expressed:

$$\delta = \frac{\sum_{j=1}^{M} \frac{c_j}{\sigma_j^2} (x_j - X_j)}{\sum_{j=1}^{M} \frac{c_j^2}{\sigma_j^2}} ,$$

where

$$c_j = \sum_{k=1}^{N} f_{jk} \, \Delta \rho_k \quad , \quad X_j = \sum_{k=1}^{N} f_{jk} \rho_k^n . \tag{5}$$

This then completes the solution for one iteration. To summarize, one computes $\Delta\rho_i$ (i = 1, N) using eq. (3), where the ρ_k^n are the values assumed at the start of the iteration. One then determines a damping factor, δ, from eq. (5), where the ρ_k^n are again the densities at the start of the iteration and the $\Delta\rho_k$ are the just-calculated adjustments. Finally, one adjusts all densities according to eq. (4).

We postpone discussion of the convergence of the iterations and briefly address the question of starting values.

3.2 Starting Values

One needs a starting value for the initial iteration. We have tried two approaches. The first was to obtain the "exact solution," described above, for a grid sufficiently coarse that the problem was tractable in the computer. The resulting densities were then projected onto the finer grid used for iterations and these values were used as starting values. The second approach was to assume a uniform density throughout of some "reasonable" value.

Both methods were acceptable, leading to solutions which converge quite rapidly. There seemed no reason to prefer the former, more elaborate method and we do not recommend it. It is slightly advantageous to select the uniform density value to give the correct value for the average sum of a set of measurements at one angle of view (i.e., to have the right "weight").

4. RESULTS

The relaxation technique has been explored in two ways. First, on computer-simulated measurements performed under a variety of conditions. Second, on a number of actual measurements made on phantom objects.

4.1 Computer Simulation

We have examined a number of different "objects" upon which measurements have been simulated by computer and then analyzed without further reference to the originating density distribution. Random errors are introduced into the measurements and effects of beam width are taken into account. In every instance so far considered the computation has converged onto the "correct" solution. That is not to say that there are not minor artifacts, but in no case has any substantial feature been observed in the analysis which was not present in the generating object, nor have significant features been missed in reconstruction.

By way of illustration we show in fig. 3 one example suggested by the paper of Cormack[2]). Fig. 3(a) shows a cross section of the phantom used in ref. [2]), and in fig. 3(b) we show the computer simulation of the phantom. "Measurements" were then made with a scan of 51 transversely separated lines of view at 40 uniformly spaced angles. A random error (standard deviation) of slightly less than 1% was introduced into all measurements. They were then analyzed on a 30×30 grid and the results (after the fifteenth iteration) are displayed in 3(c). In these plots the density is proportional to the density of displayed dots.

In fig. 3(d) we take advantage of our knowledge of the initiating density distribution to display the difference between the analysis and the true answer. In this display a horizontal slash represents a density deficit and a vertical slash a density surfeit. The density of slashes is plotted to the same absolute scale as the dot density in 3(b) and 3(c). From this representation one feature of the unfolding process is very clear, namely, that there is a strong tendency to be in error equally

above and below the true values and this type of "oscillation" about the true value occurs mainly where sharp density variations occur. Very sharp edges are not resolvable and introduce these oscillations which, nevertheless, tend to give a null contribution to any calculation of a line integral through part or all of the object.

One further feature merits attention. The density in the corners (as elsewhere) is the result of computation. Since one has a circular object and a square outline to the grid these regions might reasonably be forced to their known (near-zero) density. We do not do this, however, because we feel it is quite useful to explore regions, such as corners, of known density as a check on the success of the analysis and to offer some empirical measure of the scale of artifacts. Having said this we must add a qualifying remark concerning our treatment of negative densities. In principle these need cause no alarm. They have always been small in our experience. However, we do give them special treatment since they are unphysical and this knowledge may reasonably be incorporated. Where a cell is assigned a negative density at the end of any iteration we reset it to zero and reassign the negative density to all neighboring positive density cells in proportion to their density values. This treatment tends to "clean-up" the corners (and other areas of near-zero density) at the expense of insight into artifacts.

4.2 Measurements

No original measurements have been made by this author. However, the generosity of others in making their measurements available for analysis has been considerable. Three such sets of data have been scanned and, since they all exhibit rather different features, they will

be briefly described. One feature common to all the measurements analyzed is the comparative coarseness of the scans. This evolves naturally from the extreme tedium of making the large number of precise measurements necessary for good resolution. It is quite evident that automated data accumulation is required to realize good resolution.

The first set of data[4]) were taken on the phantom schematically represented in fig. 4(a). This object was placed in an 840 MeV alpha beam at the 184 inch Lawrence Berkeley Laboratory cyclotron, and measurements of the transmitted beam energy were made as the phantom was both translated and rotated in the beam line. The projected stopping power along the beam line was computed from the average energy degradation. The data were then analyzed to retrieve the local stopping power through a cross section of the phantom. Fig. 4(b) shows the result of that analysis. The scan involved 41 translations and 19 different view angles. The oblong analysis grid was divided into 12×24 cells. The coarseness of the scan leads to limitations on the spatial resolution which should be of the order of one-half the size of the central "spine." Such details as a slight asymmetry of the outer phantom wall and the fact that one "lung" was closer to the side of the phantom than was the other are faithfully reproduced in this reconstruction.

The second set of data[5]) were x-ray transmission data taken on a spherically symmetric annulus depicted in fig. 5(a). Since the phantom had this symmetry, only one view angle was adopted and 15 measurements were made in equal steps from the center to just beyond the outside edge. To simulate a full scan the same data were repeated as though taken at 30 different angles and these were analyzed on an 18×18

grid. The results are presented in fig. 5(b). This way of analyzing the data was adopted for convenience and is clearly not the optimum approach.

The third set of data were those reported in ref. [2]) on the phantom which is depicted in fig. 3(a). The probe was a collimated gamma-ray beam. Transmission was measured at 25 view angles and along 19 transversely spaced lines. The transverse lines were not uniformly spaced, however. They were more closely spaced towards the edge of the phantom. This feature is related to the reconstruction technique presented in ref. [2]), but we mention it here to illustrate the flexibility of our approach, which in no way depends on the way in which measurements are made (although, of course, one must know the paths along which measurements are made and the reconstruction accuracy will depend on the sampling used). We show, in fig. 6, the result of the reconstruction. It was made on a rectangular grid of uniform spacing — not the most appropriate in view of the non-uniform measurement grid. This explains why the resolution towards the edge is inferior to that achieved in ref. [2]).

5. CONVERGENCE

There are three questions which one might ask concerning the iterations. Do the iterations converge? Do they converge to a unique (and correct) solution? And, finally, do they give a reconstruction which predicts reasonable values for the observations?

In fig. 7 we plot the sum of squares (as defined in eq. (2)) against the iteration number for the analysis of Carmack's data[2]), the results of which were presented in fig. 6. No example has been encountered in which such convergence did not obtain. We have no

proof, however, that such convergence will necessarily occur and can only state that, in our experience, it always has.

Similarly, in regard to the uniqueness of the situation, we have been unable to prove a uniqueness theorem. One might worry, as some have[1,3]), that ambiguous solutions may exist. In some cases, notably when a very limited number of measurements are made, there are certainly very real ambiguities inherent in the measurements. This is the case with the example presented in ref.[3]). One must realize that in such cases the ambiguity is real and any method of analysis must be responsive to that problem which is a consequence of the inadequacy of the measurements. The ability to converge on the correct solution is directly related to the high degree of redundancy in the measurements. One must make substantially more measurements than the number of pieces of information one hopes to extract (see below). If that is done, it is our experience that the iterative relaxation technique does converge to the correct solution even when substantially different (and unreasonable) starting values are used.

Finally, one must note that the convergence may be to rather poor values of the least-squares parameter. Fig. 7, for example, shows a convergence of the sum of squares to a value of about 200 per degree of freedom, which is quite enormous compared to the value of unity expected if the problem was dominated by random statistical errors[6]). One may understand this in terms of two effects. The first and more serious problem has to do with the computational premise that the object may be represented by an array of cells with density uniform within each cell. Clearly, if the scanned object has features whose density varies rapidly

over the dimension of a cell side the representation cannot be adequate. In this connection alternative assumptions might be made. One might assign densities on a grid and interpolate between points using any variety of schemes. These methods are equivalent to forcing various degrees of smoothness on the reconstruction. They may well be called for in some instances; certainly the relaxation technique can trivially accommodate such a procedure.

The second effect leading to poor values of the sum of squares has to do with the nature of the probe. Beam scatter and uncertainties in the beam profile will introduce errors in the reconstruction which will lead to poor least-squares values.

6. RESOLUTION

6.1 Choice of Number of Measurements

Consider a specific grid of, say, $n \times n$ squares. What is the maximum number of measurements which can give useful information? Clearly, two measurements which are so closely spaced that they pass through almost the same cells with almost the same average path length in each cell will not yield substantially different information from each other. Very crudely one might say that a translation of one-half of a cell side is needed to produce substantially different information. This would lead to $2n$ measurements at each angle of view.

The maximum number of useful view angles may be estimated as suggested in fig. 8. The smallest useful included angle between measurements common to a cell at one edge of the grid is that which leads to a separation of one cell width on the other side of the object. This is an angle of $1/n$ radians and leads to a maximum number of view angles of $2\pi n$.

Thus, these simple considerations suggest that $4\pi n^2$ measurements are needed to resolve n^2 cells. In practice we have found this to be an overestimate. We have found $3n^2$ to be adequate. We present this intuitive estimate less as a hard and fast guide to estimating the number of measurements required for a given resolution than for the insight it offers in understanding the planning of measurements. For example, it is clear than one cannot get the same information out of a fixed number of measurements by increasing the number of transversely separated projections at the expense of the number of angles of view (or vice versa). The number of scan angles must be quite large for reasonable resolutions (an experimentally disconcerting requirement).

6.2 Measurement Accuracy

The accuracy with which measurements must be made clearly depends on the density resolution required. If a region of $k \times k$ cells has density $\rho_0 + \Delta\rho$ which must be distinguished from a background density ρ_0 then measurements through that region will differ by a fraction $(k/n)\,(\Delta\rho/\rho_0)$ from other measurements (the entire object is assumed to be divided into $n \times n$ cells). If the average measurement value is $\bar{X}$ with standard deviation σ the requirement

$$\frac{\bar{X}}{\sigma} \; \frac{\Delta\rho}{\rho_0} \; \frac{k}{n} \gtrsim 1 \qquad (6)$$

is clearly a sufficient condition for observing the interesting region. (Note, however, that there is a less stringent requirement, discussed below).

There is another aspect to the measurement accuracy which is harder to quantify. It emerges from the observation that one does

not improve the reconstruction without limit as one increases the measurement accuracy. This is because the representation of the object as a set of discrete cells with uniform density (or any other interpolative representation) is imperfect. Qualitatively one might say that the difference between a measurement on the true object and the measurement which would result from the optimum $n \times n$ cell approximation of the object gives a measure of the level of accuracy which cannot usefully be exceeded. What that level is will depend greatly on the structure of the scanned object.

7. RESTRICTED RANGE OF SCAN ANGLES

Ideally one would wish to make measurements in the full range of possible angles, namely 180°, and this was done for all the preceding examples. However, it may not be possible to make the full scan in practice. Restriction of the range of scan angles does severely limit the resolution achievable. This matter was quantitatively discussed in ref. [3]). We do not have a quantitative estimate of this effect in terms of the iterative relaxation technique. One can see intuitively that if one confines oneself to ±45° about the y-axis (as in fig. 9a) one will have less information (hence, resolution) about structure parallel to the y-axis than that parallel to the x-axis. In figures 9b, c, and d we show the reconstruction of the object depicted in fig. 9a from measurements made respectively in the range ±45°, $\pm 67\frac{1}{2}°$, and the full ±90°. In all three cases the same 20×20 cell grid is used and the measurements comprised 51 translations $\times 20$ angles. The degradation of resolution in the y direction relative to the x direction is quite striking.

8. ARTIFACTS AND THE LIMITS OF RESOLUTION

The central problem when faced with a reconstruction is to be able to say whether some selected feature is "real" or on artifact. One might ask the question, "If I remove the structure and replace it by the density of the local background is the modified reconstruction significantly less able to fit my measurements?" One should also, perhaps, require that all cells be uniformly renormalized after removal of the structure so that the "weight" of the reconstruction remains unaltered.

Implicit in our previous discussions has been the fact that one has two separate resolutions to deal with: spatial and density resolution. The abilities to resolve spatial detail and to detect density variations are distinct. They are, however, clearly correlated. One might expect that a very small object could more readily be detected if it was of a very different density from its surround. We will now suggest a quantitative measure of this effect.

The question posed above has a direct answer from the theory of least-squares fitting. The standard deviation in the value of a parameter (such as the density of some feature) is estimated by the change in the parameter necessary to increase the sum-of-squares parameter by unity. In any given situation one may determine that change by direct computation. Here we examine the general case. Consider an object of $n \times n$ cells all of which have density ρ except for a clump of $k \times k$ cells somewhere in the object which have density $\rho + \Delta\rho$. We then estimate the change in $\mathcal{M}^2$ when we set the $k \times k$ cells back to ρ (and then increase all cells by $(k/n)^2 \Delta\rho$ to maintain the same weight). We require this

change to be greater than or of the order of unity for the clump to be observable. This leads to the requirement

$$\frac{\bar{X}}{\sigma} \frac{\Delta_\rho}{\rho} \left(\frac{k}{n}\right)^{3/2} \sqrt{M} \gtrsim 1. \qquad (7)$$

Here M is the total number of measurements, $\bar{X}$ is the average value of a measurement, and σ is the average measurement error.

This formula should be treated more as a suggestive estimate than a hard and fast quantitative resolution limit. For one thing the prescription to raise m^2 by unity from its minimum value is only applicable when the minimum value is reasonable (within a few $\times \sqrt{D}$ of the number of degrees of freedom, D). However, as we have previously pointed out, this may not be the case in practice. One finds oneself abandoned by statistical theory at that point. One tactic of desperation is to readjust the estimates of error, σ, by the amount which will force the minimum value of m^2 to be equal to the number of degrees of freedom.

Eq. (7) is interesting in that it suggests that the critical dimension for resolution of some feature is something intermediate between its diameter and area. In any event, the parameter which must be used to characterize the spatial resolution of any reconstruction is the ratio of the size (diameter or area) of the feature to that of the entire object.

To give concreteness to eq. (7) we give a numerical example. Suppose that one makes 10 000 measurements, each of 3% accuracy, and asks how large a feature must be to be distinguishable if its density is 3% different from its background. In this case k/n is about 1/22, which means that the diameter of the feature must be at least one

twenty-second that of the scanned field. This is, of course, rather a modest spatial resolution but it is characteristic of this kind of reconstruction.

9. COMPARISON WITH OTHER TECHNIQUES

There are two general classes of analysis technique. On the one hand there are those methods which, as with that presented here, attempt to generate a fully unfolded distribution, disentangling the contributions of separate cells. On the other hand there are tomographic techniques which seek to accentuate specific regions within the viewed object by defocusing contributions from all other regions. The tomographic techniques offer especially simple methods of analysis which can often be implemented by purely mechanical programming of the measuring apparatus. They are subject to the serious flaw of always superposing a background of (imperfectly) defocused structures on the region examined. We do not further consider the tomographic analysis here.

In regard to the techniques which attempt a full unfolding it should be noted that the presently available analyses are all comparable in the amount of information which they extract from a given set of measurements. The choice of technique must therefore lie in considerations of simplicity, feasibility in terms of available computational capacity, and applicability in relation to the particular problem on hand.

Two alternative techniques have been published[7]. In the first[1,2], the two-dimensional problem is separated into a set of one-dimensional integral equations of a function with solely radial variation. The measurements are expanded in a sine series with coefficients identical to

those of the radial density function when expanded in a limited series of Zernicke polynomials.

The second technique[3]) depends on the observation that the Fourier transform of a projected view is just the value which the three-dimensional Fourier transform takes on a plane through the origin in Fourier space. Projections at different angles build up the Fourier transform on different planes and enable one to construct the full three-dimensional transform, which can then be inverted to regain the three-dimensional reconstruction in real space. In the inversion one must interpolate between measured values of the transform, since sampling points do not generally coincide with measured values and this interpolation can introduce artifacts. Alternative interpolation procedures have been investigated[8]). The method of Cormack[1,2]) is mathematically equivalent to the Fourier transform technique[3]) but, clearly, differs substantially from it in practical application.

We suggest now some features unique to the iterative relaxation technique which might make it advantageous in certain circumstances.

9.1 Versatile Scanning

As has been emphasized already, this method can accommodate any series of measurements, since the only way the geometry enters into the computation is through the f_{jk} of eq. (1). Thus analysis of x-ray projections taken with short-focal-length setups (and consequently with highly divergent beams) presents no problem. Similarly, omitted or repeated measurements and irregularly spaced measurements are easily handled.

Both the beam profile at the entrance to the scanned object and the variation of the profile with depth in the object are features which enter directly into the calculation of the f_{jk} and can consequently be fully accounted for.

9.2 Versatile Reconstruction Grid

While the discussion of the method has referred to a Cartesian grid of cells there is no reason at all to make that restriction. Clearly one can employ polar-coordinate grids or other regular arrays, perhaps matching the cell size in any given region to the expected structure there. Indeed, this can be extended to the much more general situation in which the scanned object is formed of a large number of regions of complex geometric form corresponding to the known configuration of the object. One could then allow the density of each region to be varied to fit the measurements. Having suggested this, one might proceed to the logical conclusion and attempt not only to modify the density of each region but also its boundaries. To do this one would have to recalculate the f_{jk} after each iteration, but that is not an unduly burdensome task. If each region were represented by an octagon one would have seventeen parameters (a density and eight coordinate pairs) for each octogon. With 10 000 measurements one could readily analyze situations having as many as 100 such variable regions.

9.3 Extra Parameters

We have implicitly assumed that each cell contributes to the measurement through a single parameter (such as its density or absorption coefficient or stopping power). One might easily imagine associating with each cell additional parameters such as, for example, the effective

atomic number of the cell material. One would then search in a multi-dimensional parameter space for the best fit to the measurements. Provided the measurements were sensitive to all the parameters, one could hope to reconstruct them all simultaneously.

9.4 Sharp Edges

We were initially led to this particular reconstruction technique by the desire to simulate the internal structure of the human body. It seemed reasonable to look for a representation which was capable of simulating large areas of fairly uniform density bounded by extremely rapid density variations such as one sees at a bone-muscle interface. The spatial resolution implied by the cell size has its counterpart in the highest-frequency component used in, say, the Fourier transform technique. However, there is a slight difference in that the use of a high frequency to effect a rapid density variation at one boundary results in the presence of high-frequency components everywhere in the reconstruction. These can be largely cancelled out everywhere else, but there is always a residual high-frequency component which leads to the typical oscillatory character of such reconstructions. In the iterative relaxation technique any cell may differ in density by any amount from its neighbors without forcing the same high-frequency response elsewhere in the system.

10. CONCLUSIONS

We have developed an iterative relaxation technique for resurrecting a three-dimensional distribution from a series of two-dimensional projections derived from it. The technique has been applied successfully to a number of computer simulations and actual laboratory measurements.

The ability to retrieve the full three-dimensional originating distribution is now well established. The question of the most appropriate technique remains open and the answer will depend on details of the problem. We have presented some advantageous features of the iterative relaxation technique.

We discuss the achievable resolution. A typical situation, which is reasonably undemanding in terms of computational capacity, would be the analysis of about 10 000 measurements to generate densities on a 50×50 element grid.

ACKNOWLEDGMENTS

I am grateful to Dr. C. Tobias for bringing this problem to my attention and for several subsequent discussions. Drs. D. Chesler, A. M. Cormack, and J. Lyman have made the raw data from their (separate) measurements available to me and I have benefitted from discussions with all of them. This work was largely supported through the Donner Laboratory.

Footnote and References

[1]) A. M. Cormack, J. Appl. Phys. 34 (1963) 2722.

[2]) A. M. Cormack, J. Appl. Phys. 35 (1964) 2908 and private communication. We are grateful to Dr. Cormack for supplying us with the raw data for the experiment described in this reference.

[3]) R. A. Crowther, D. J. DeRosier, and A. Klug, Proc. Roy. Soc. (London) A317 (1970) 319.

[4]) These data and much helpful discussion were kindly provided by Dr. J. T. Lyman of Lawrence Berkeley Laboratory.

[5]) We are indebted to Dr. D. Chesler for these data.

[6]) The value of the sum of squares parameter is exceptionally high in this case. We believe the poor fit is the result of the poor match between the uniform analysis grid and the nonuniform measurement grid.

[7]) A third analysis technique has recently been developed by D. Chesler (private communication). This technique may be particularly well suited to analysis with small computers.

[8]) O. Tretiak, private communication.

Figure Captions

Fig. 1. Basic series of measurements comprising a series of projections along parallel transversely separated paths, each series being made at a number of different orientations. Each arrowed line represents the path along which one measurement is made.

Fig. 2. Reanalysis grid with one typical measurement schematically indicated.

Fig. 3. Computer simulation of Cormack's experiment (ref. [2]).

(a) Original phantom

(b) Computer simulation of phantom. Density is linearly proportional to the dot density.

(c) Computer reconstruction of phantom

(d) Difference between reconstructed and true densities. Horizontal slashes are deficits and vertical slashes are density excesses.

Fig. 4. Lyman experiment using 840 MeV alpha particles (ref. [4]).

(a) Schematic representation of phantom

(b) Reconstruction.

Fig. 5. Chesler experiment (ref. [5]).

(a) Phantom

(b) Reconstruction.

Fig. 6. Reconstruction of phantom shown in fig. 3a based on analysis of the original Cormack data (ref. [2]).

Fig. 7. Convergence for the Cormack experiment. The sum of squares per degree of freedom is plotted against iteration number.

Fig. 8. Least significant angular interval between measurements for a given level of resolution (grid size).

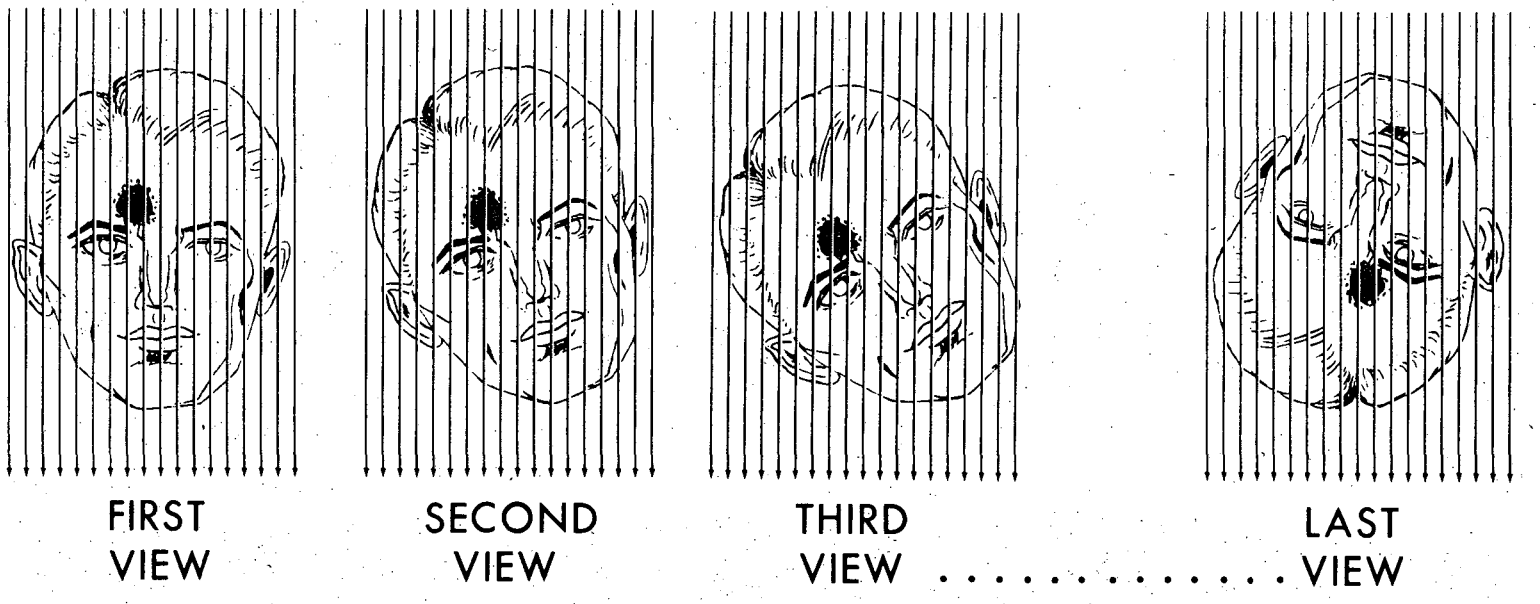
Fig. 9. Computer simulation of measurements made with a restricted set of scan angles.

(a) Original phantom with indication of central axis relative to which angles were measured

(b) Reconstruction from measurements made between -45° to +45°

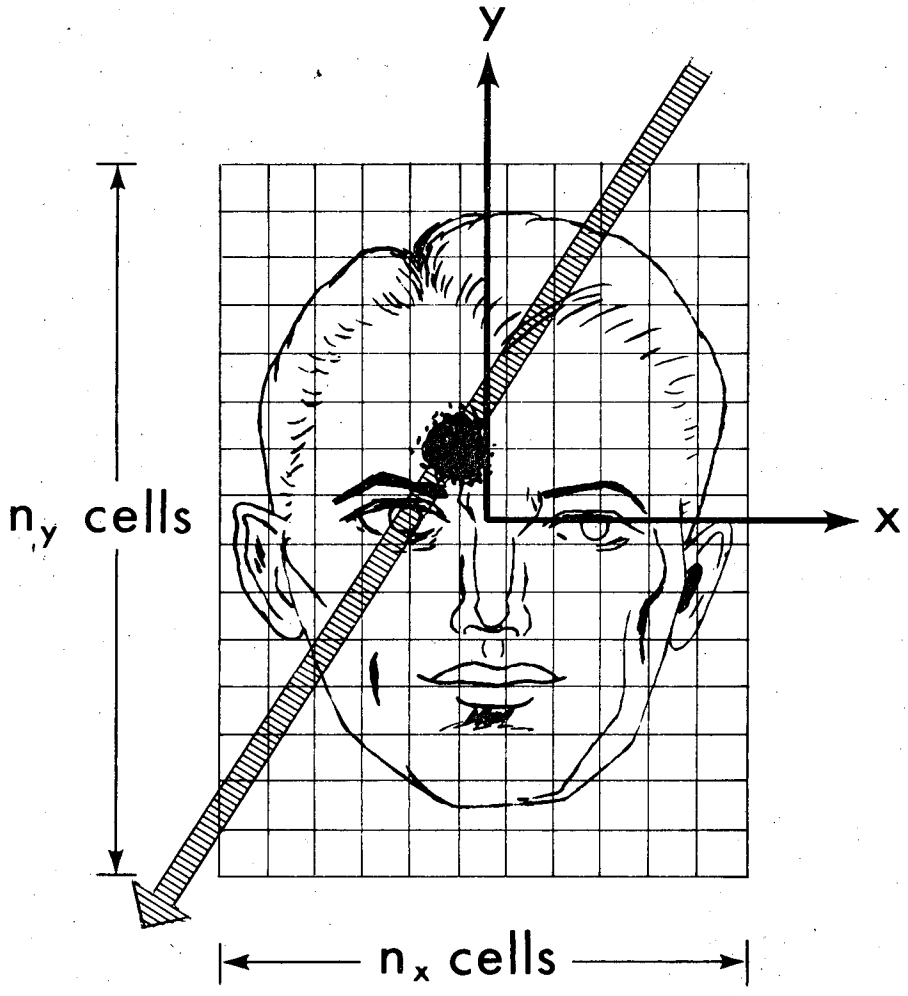
(c) Reconstruction from measurements made between -67.5° to +67.5°

(d) Reconstruction from measurements made between -90° to +90° (full range of angles).



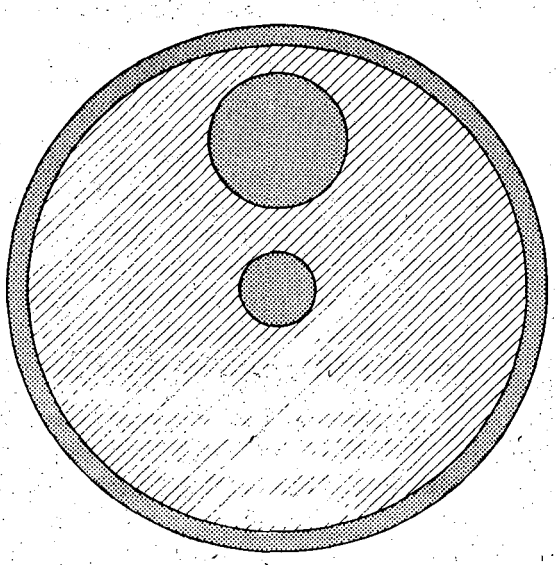
DBL 7112 6115

Fig. 1

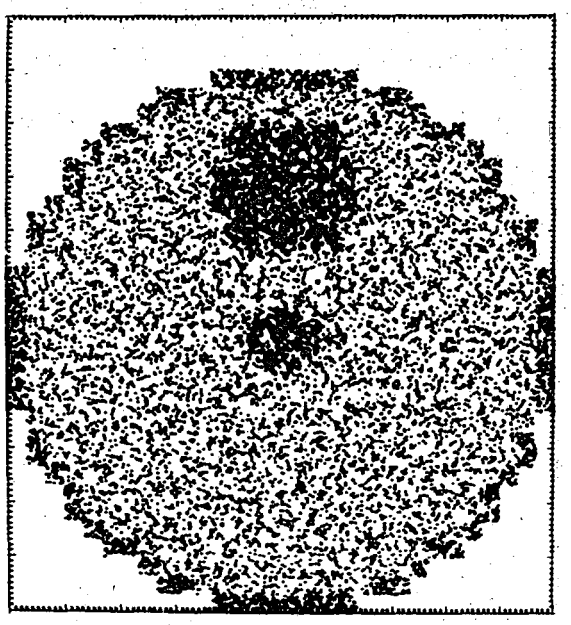


DBL 7112 6117

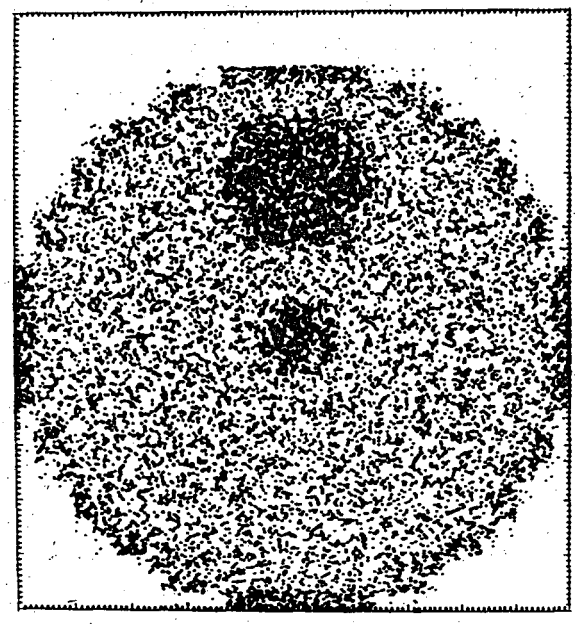
Fig. 2



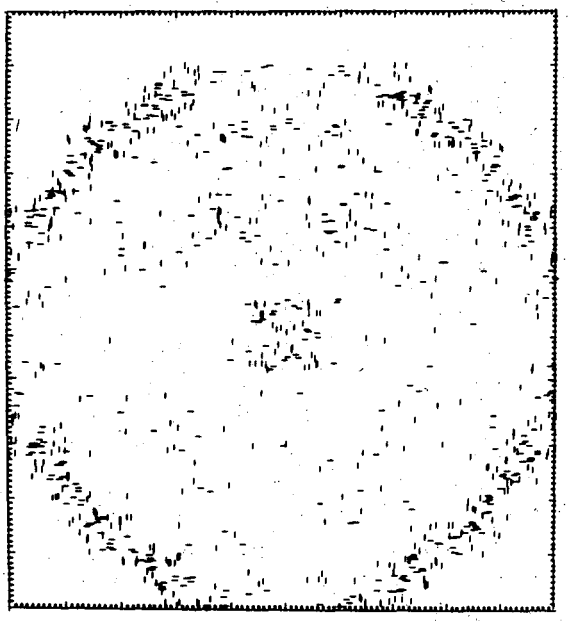
(a) ORIGINAL PHANTOM



(b) SIMULATED PHANTOM



(c) RECONSTRUCTION



(d) ERROR

DBL 7112 6120

Fig. 3

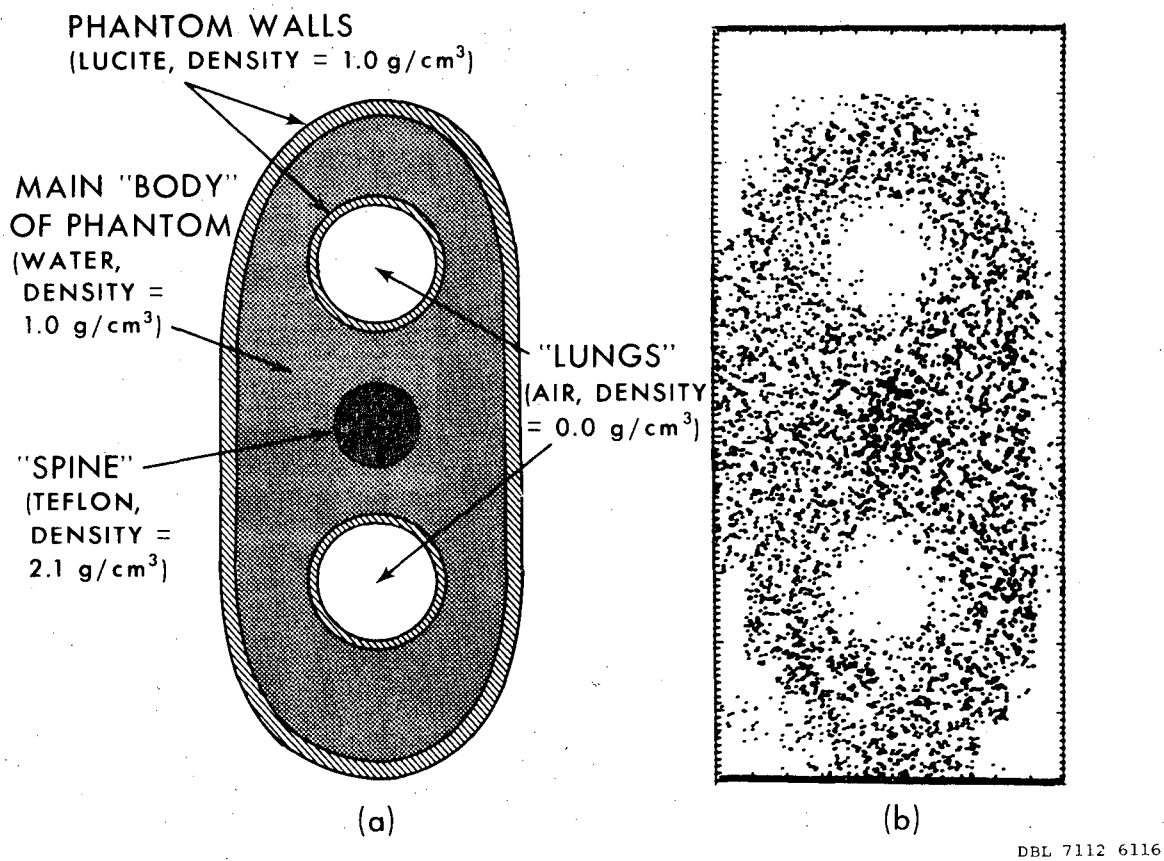
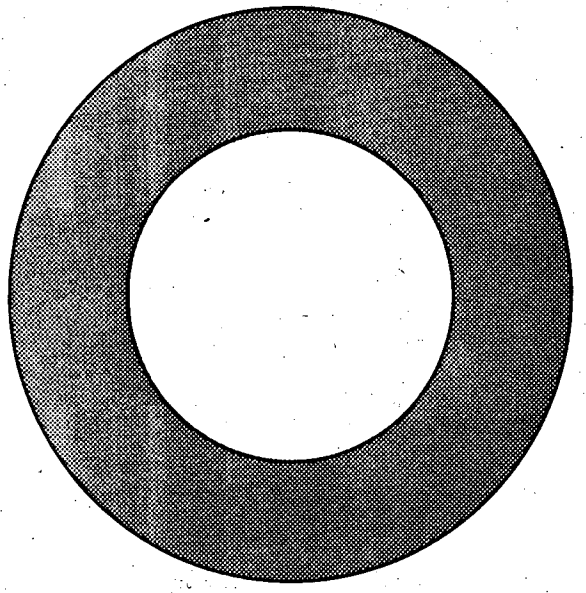
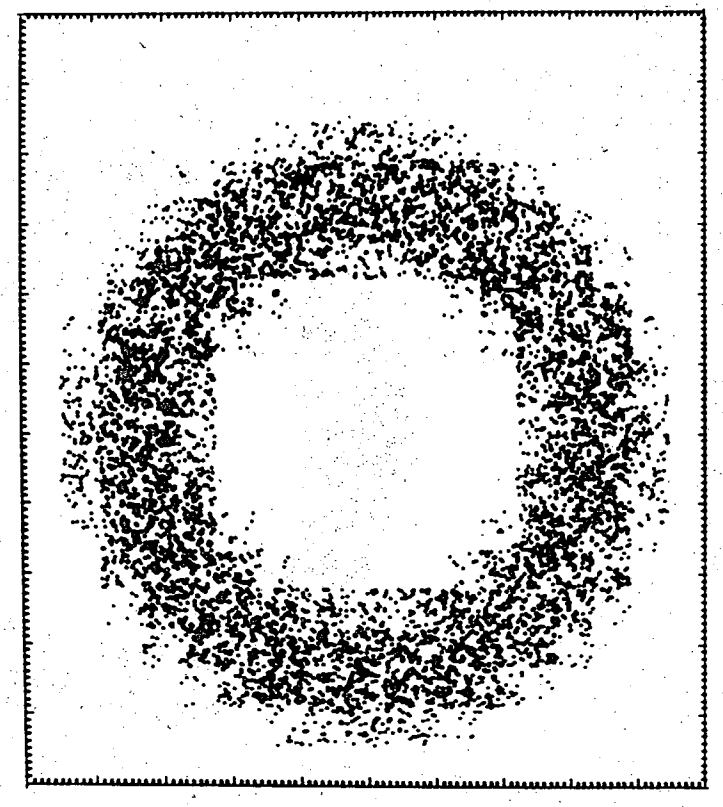


Fig. 4



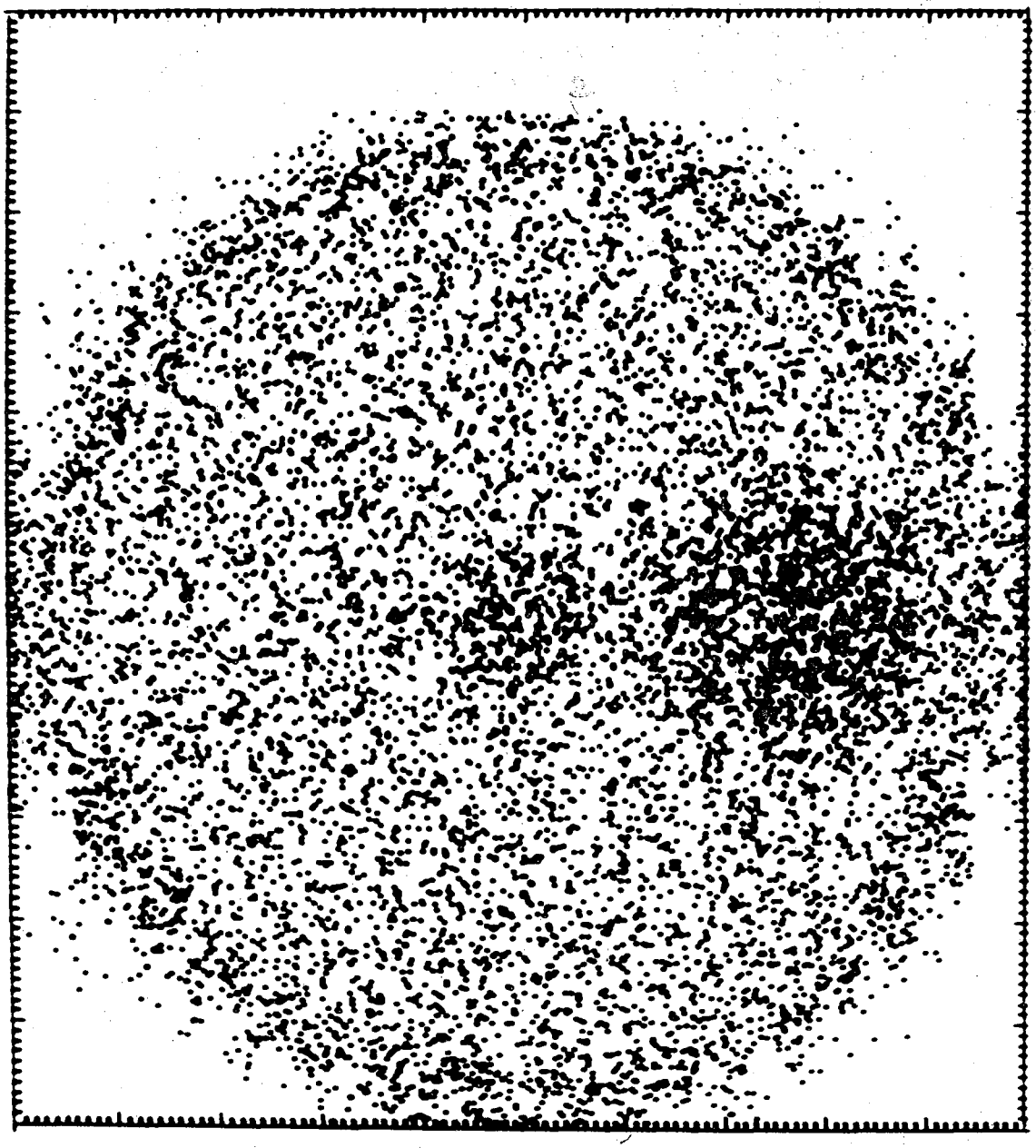
(a) PHANTOM



(b) RECONSTRUCTION

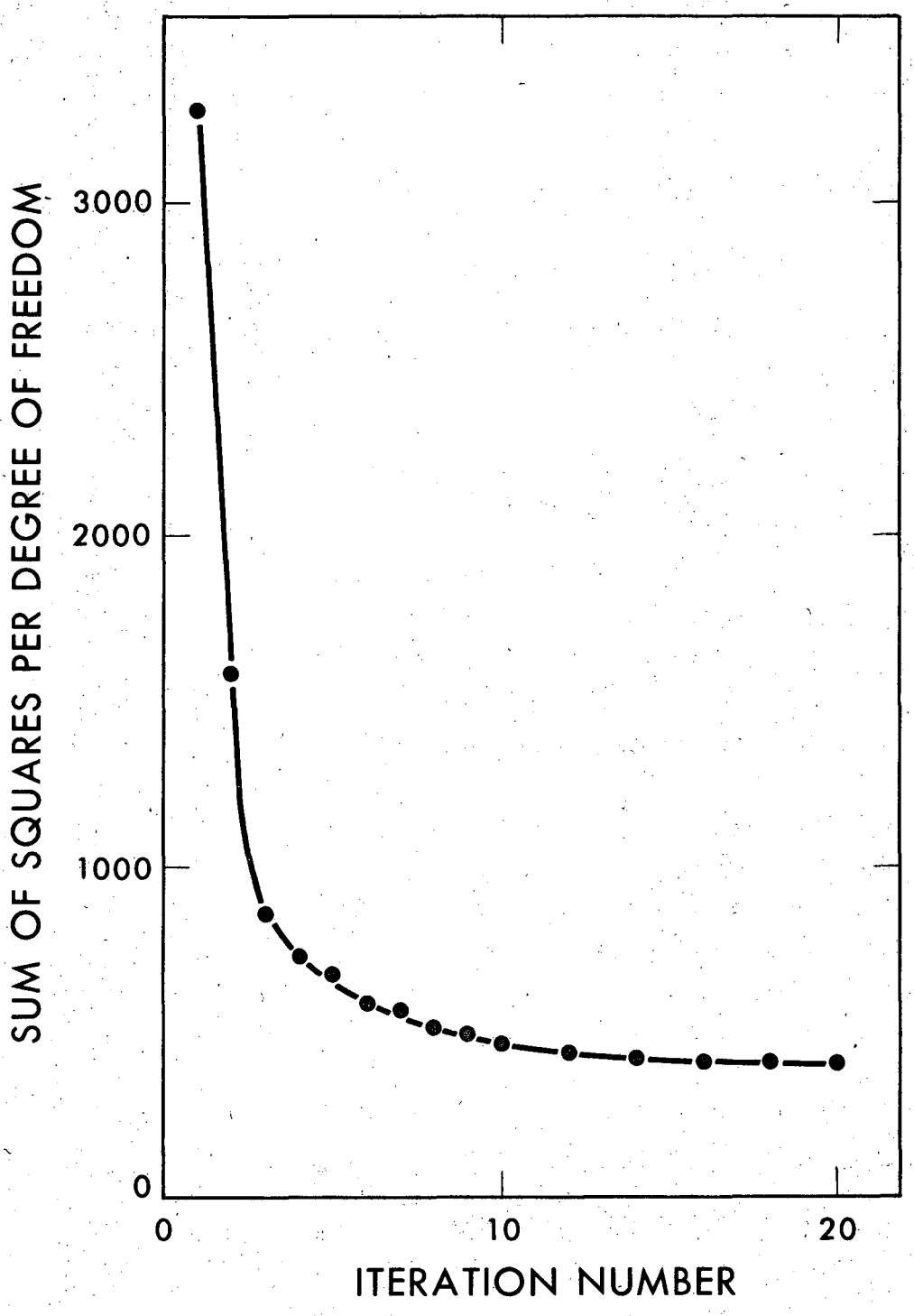
DBL 7112 6119

Fig. 5



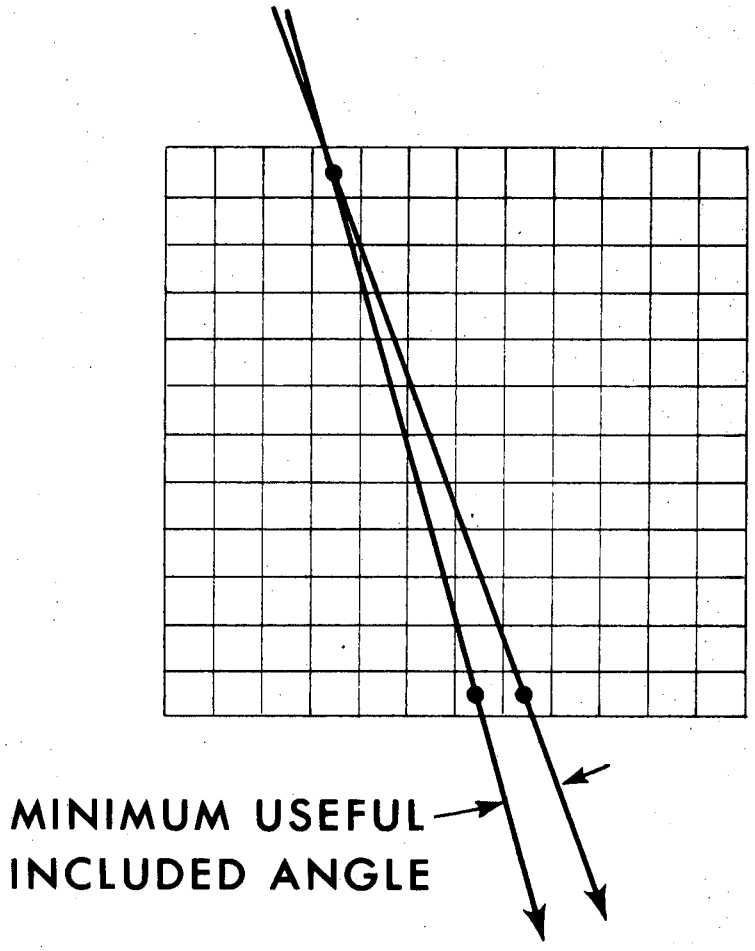
DBL 7112 6118

Fig. 6



DBL 7112 6113

Fig. 7



DBL 7112 6114

Fig. 8

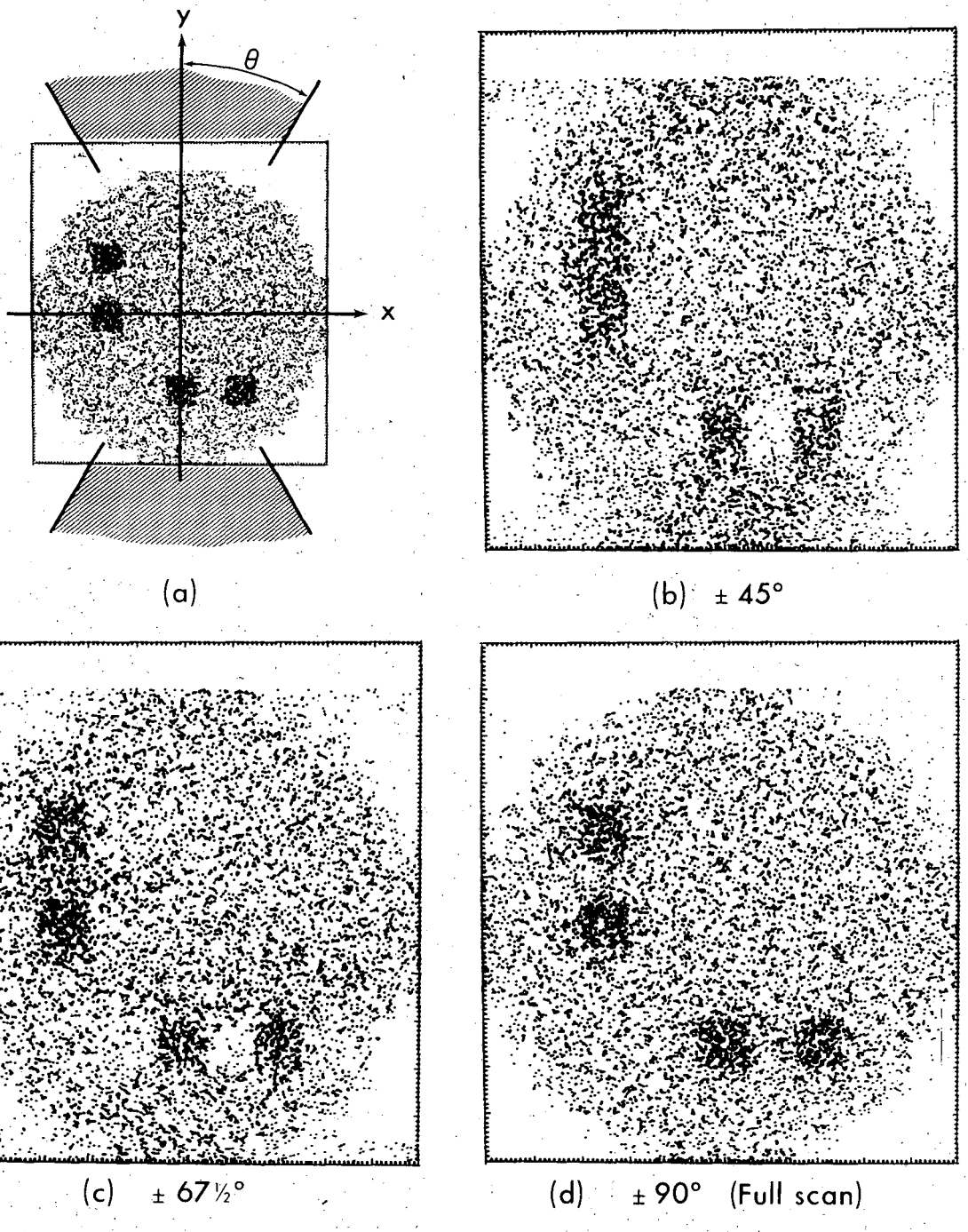


Fig. 9

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