LA-6894-MS

Informal Report

UC-28 and UC-32

Issued: August 1977

MASTER

Linac Beam Core Modeling from Wire-Scanner Data

Alan G. Law*

*Visiting Staff Member. University of Regina, Department of Computer Science, Regina, Saskatchewan, Canada S4S 0A2.

los alamos
scientific laboratory
of the University of California
LOS ALAMOS, NEW MEXICO 87545

An Affirmative Action/Equal Opportunity Employer

UNITED STATES
ENERGY RESEARCH AND DEVELOPMENT ADMINISTRATION
CONTRACT W-7405-ENG. 36

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Printed in the United States of America. Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161
Price: Printed Copy ~~\$3.50~~ Microfiche \$3.00

4. 23
This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Energy Research and Development Administration, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or 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.

LINAC BEAM CORE MODELING FROM WIRE-SCANNER DATA

by

Alan G. Law

ABSTRACT

NOTICE
This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Energy Research and Development Administration, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or 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.

This study introduces mathematical modeling of accelerator beams from data collected by wire scanners. Details about a beam core $D(x, x', y, y')$ are examined in several situations: (a) for a discretization of the projection into xy -space, a maximum-entropy solution and a minimum-norm solution are developed and discussed, (b) for undiscretized xy -subspace, a two-dimensional Gaussian approximation $D(x, \cdot, y, \cdot) = a \exp\{\alpha(x-x_0)^2 + \beta(x-x_0)(y-y_0) + \gamma(y-y_0)^2\}$ is obtained by least squares, and (c) for four-dimensional space, the fit of a single Gaussian to data from a succession of wire scanners is investigated.

I. INTRODUCTION

The Los Alamos Meson Physics Facility (LAMPF) is a national facility which is centered on a linear accelerator designed to provide a proton beam with energy up to 800 million electron-volts and an average current of one milliamper. The accelerating structure, approximately one kilometer in length, includes a variety of measuring devices for continual monitoring and control of the beam. One type of device is the wire scanner¹--two separate wires, usually of carbon or tungsten and a few mils thick, are driven across the beam pipe and the signal on each wire is read at each step. Typically, the number, n , of steps selected for a 3-cm path will be between 40 and 100 (see Fig. 1). At any position, the current (after amplification then conversion to a voltage signal before the final digital form) gives the total number of protons hitting all along the wire; thus, these measurements do not yield an estimate of the number of particles at a particular point on the wire. Such finer detail would be useful, for instance, in determining if the beam has a center and where it is relative to the axis of the pipe,

or if there are the undesirable "halo" components in addition to a single central "core." If the situation is to be expressed by some unknown density function $D(x,y)$ whose value at (x,y) should specify the number of protons impinging there at that instant, then information available to the modeling process is very limited. More specifically, for $2n$ lines, L , prescribed in the domain (n of them being parallel to the x -axis and the other n being parallel to y), the total number of particles colliding with each is known, - in analytic terms, each of the $2n$ line integrals

$$\int_L D(x,y) ds$$

should be a specified value. This problem of reconstructing a function of two variables from a subset of its projections receives much attention²⁻⁵ in areas such as radio astronomy, electron microscopy, medical radiology and holographic interferometry, and both the angles and number of directions available for viewing are of fundamental importance. Unfortunately, a single wire scan presents only two views. In fact, the situation is even worse than this: the actual beam is a function of six or four variables according as it is bunched or not, so a wire-scan density function $D(x,y)$ would simply be one projection obtained through multiple integration of an unknown beam density function of several variables. Thus, even if the effects of noise in the signals could be ignored, the data from a single wire scan are insufficient for determination of merely the projection into xy -phase space, let alone of the multidimensional beam itself. However, since linear approximations of the longitudinal transformations between adjacent monitoring stations are known for drifting, accelerating, and quadrupole effects, a succession of scans provides data about several additional projections of the beam into the xy -subspace. A second useful approach results from imposing some selected structure on the beam; for example, a Gaussian form is often hypothesized for a projection emerging in one of the phase spaces^{6,7}.

In this report, a maximum-entropy form, a minimum-norm form and a single-Gaussian form are obtained for data from a single wire scan and then the fitting of a four-variable Gaussian to consecutive scans is developed. It should be emphasized that this development is restricted to the setting of a single core for the beam. As operating currents are raised and fringe

effects become increasingly important, more complex modeling, particularly for halo, will be needed as an aid to firmer understanding of particle beam characteristics.

II. SINGLE WIRE-SCAN RECONSTRUCTION--DISCRETE CORE MODELS

A. Formulation

Data are collected from a single pass of one wire scanner as it crosses the beam by distinct steps. In attempting to model the view of the beam at this cross section, two different approaches may be employed: (1) the xy-beam is to be described by some continuous function, $D(x,y)$, whose domain is the entire cross section, but which is sampled at each step, or (2) the cross section is partitioned into "bins" and the number of particles to be assigned to each bin is sought. The former approach is the one taken in Sec. III.

Let $x_1 < x_2 < \dots < x_n$ and $y_1 < y_2 < \dots < y_n$ form a grid of points (x_j, y_i) over a region of the xy plane and consider an (unknown) nth-order matrix $M = (M_{ij})$ whose ijth entry represents the number of particles to be associated with the node (x_j, y_i) . With the measured row-sum data R_i and column-sum data C_j , the criteria imposed on M are simply that

$$\left. \begin{aligned} \sum_{j=1}^n M_{ij} &= R_i, \quad 1 \leq i \leq n \\ \sum_{i=1}^n M_{ij} &= C_j, \quad 1 \leq j \leq n \end{aligned} \right\} \quad (1)$$

For a solution matrix M to exist, it is necessary that $\sum_{i=1}^n R_i = \sum_{j=1}^n C_j$;

the proof of Lemma 1 shows that this condition is also sufficient for existence.

Now, let $T \equiv \sum_{i=1}^n R_i - \sum_{j=1}^n C_j$: if $T \neq 0$ (because of noise in data collection or transmission, for instance), the data can be adjusted by replacing R_i with $R_i - \frac{T}{2n}$ and C_j with $C_j + \frac{T}{2n}$. It will therefore be assumed that the observed data used always satisfy the

Consistency Condition

$$\sum_{i=1}^n R_i = \sum_{j=1}^n C_j; \quad (2)$$

so that

$$S = \sum_{i=1}^n R_i = \sum_{j=1}^n C_j \quad (3)$$

is well defined.

System (1) is an algebraic system of $2n$ equations, in the n^2 unknowns M_{ij} , which may be represented either by

$$\begin{array}{cccc|c} M_{11} & M_{12} & \cdots & M_{1n} & R_1 \\ M_{21} & M_{22} & \cdots & M_{2n} & R_2 \\ & \cdots & & & \vdots \\ M_{n1} & M_{n2} & \cdots & M_{nn} & R_n \\ \hline C_1 & C_2 & \cdots & C_n & S \end{array} \quad , \quad (4)$$

or by

$$HX = Z, \quad (5)$$

where $X \equiv (M_{11}, M_{12}, \dots, M_{1n}, M_{21}, M_{22}, \dots, M_{2n}, \dots, M_{n1}, M_{n2}, \dots, M_{nn})^T$,

$Z = (C_1, C_2, \dots, C_n, R_1, R_2, \dots, R_n)^T$, and H denotes the $2n \times n^2$ matrix whose blocked form is

$$\left(\begin{array}{c|c|c|c|c} I & & & & \\ \hline & I & & & \\ \hline & & \cdots & & \\ \hline & & & I & \\ \hline U_1 & U_2 & & & U_n \end{array} \right),$$

wherein I is the n th-order identity and U_k ($1 \leq k \leq n$) is the n th-order matrix having each element of its k th row equal to one and all other elements equal to zero. Sometimes Eq. (5) is more convenient to use than Eq. (4). Elementary row operations show quite readily that the last row of H is a linear combination of the other rows and also that its predecessors form a linearly independent set. Consequently, H has rank $2n-1$.

Lemma 1. Linear system (5) has infinitely many solutions.

Proof. Since the coefficient matrix has rank $2n-1$ but there are n^2 unknowns, it need only be shown that Eq. (5) has at least one solution. However, it is easy to verify directly that one solution matrix for Eq. (4) is

$$E = (E_{ij}) = \left(\frac{R_i C_j}{S} \right). \quad (6)$$

Solution (6) may be employed to show that requiring the elements of a solution matrix to be positive still does not ensure uniqueness.

Lemma 2. Suppose that each R_i is positive and each C_j is positive. Then Eq. (4) has infinitely many solutions M in which all elements are positive.

Proof. For the particular (positive) solution [Eq. (6)], let

$$\begin{aligned} \bar{E} &\equiv \min_{1 \leq i, j \leq n} E_{ij} \\ &= E_{IJ}. \end{aligned}$$

Select any ϵ in the interval $(0, \bar{E})$. Unless $\bar{E}$ occurs only at a corner element for which an obvious modification of the following scheme is required, simply replace E_{IJ} , $E_{I, J+1}$, $E_{I+1, J}$ and $E_{I+1, J+1}$ by, respectively, $E_{IJ} - \epsilon$, $E_{I, J+1} + \epsilon$, $E_{I+1, J} + \epsilon$ and $E_{I+1, J+1} - \epsilon$ to obtain another positive-termed solution to Eq. (4).

Two interesting solutions of Eq. (4) [or Eq. (5)] materialize when minimum-norm or maximum-entropy conditions are imposed.

3. The Minimum-norm Solution

Let $||\cdot||$ denote the Euclidean norm defined on n^2 -tuples; in other words, $||\cdot||$ is the norm defined on n th-order (real) matrices $A = (A_{ij})$ by

$$||A||^2 = \sum_{i=1}^n \sum_{j=1}^n A_{ij}^2. \quad (7)$$

It is well known⁸ that, for any nonsquare algebraic system [Eq. (5)], there exists a unique n^2 -tuple which, from among all those that minimize $||HX - Z||$, has minimum norm. In fact, this best n^2 -tuple is⁸ simply H^+Z , where H^+ is the pseudoinverse of H .

Lemma 3. For the $2n \times n^2$ matrix described in Eq. (5), its Moore-Penrose pseudoinverse is the $n^2 \times 2n$ matrix H^+ given by:

$$H^+ = \frac{1}{2n^2} \begin{pmatrix} D & V_1 \\ \hline & \\ D & V_2 \\ \hline & \\ & \vdots \\ \hline & \\ D & V_n \\ \hline & \end{pmatrix}, \quad (8)$$

where D is the n th-order matrix with diagonal elements equal to $2n-1$ and off-diagonal elements equal to -1 , while the n th-order V_k ($1 \leq k \leq n$) has each element of its k th column equal to $2n-1$ and all others equal to -1 .

Proof. An elementary but lengthy calculation confirms that H^+ , as given by Eq. (8), indeed satisfies necessary and sufficient criteria⁸ for being the pseudoinverse of H . That is,

- [i] HH^+ and H^+H are symmetric, and
- [ii] $HH^+H = H$, and
- [iii] $H^+HH^+ = H^+$.

In view of Lemma 1, $||HX-Z|| = 0$ has many solutions, but the one of minimum norm can now be calculated directly since H^+ and Z are known. The results may be summarized as:

Theorem 1. Let $C_1, C_2, \dots, C_n, R_1, R_2, \dots, R_n$ be given data which satisfy the Consistency Condition, and let

$$N = (N_{ij}) = \left(\frac{R_1}{n} + \frac{C_1}{n} - \frac{S}{n^2} \right). \quad (9)$$

Then:

- (a) N is a solution of Eq. (4), and
- (b) for any other solution M of Eq. (4),

$$||N|| < ||M||.$$

The minimum-norm solution N described in Eq. (9) has interesting theoretical properties; for example, this particular reconstruction procedure is seen to be linear since $H^+(a_1X_1 + a_2X_2) = a_1H^+X_1 + a_2H^+X_2$. However, the closed-form Eq. (9) indicates that N can introduce a smearing effect: if

the R_i and C_j data are generated from an n th-order matrix whose ij th element is 1 while all its other elements are zero, the corresponding reconstruction N will blur this sharpness by introducing a cross of nonzero entries into the i th row as well as the j th column. Such smearing has been observed in extensive numerical investigation, and the results from a typical test case are shown in Fig. 2.

C. The Maximum-Entropy Solution

Let $C_1, C_2, \dots, C_n, R_1, R_2, \dots, R_n$ be given consistent data and let $\mathcal{E}$ denote the entropy function* defined on n th-order, non-negative matrices $A = (A_{ij})$ by

$$\mathcal{E}(A) = - \sum_{i=1}^n \sum_{j=1}^n (A_{ij}/S) \log (A_{ij}/S) . \quad (10)$$

With the notation

$$a_{ij} \equiv A_{ij}/S ,$$

a maximum of

$$\mathcal{E}(a_{11}, \dots) = - \sum_{i,j=1}^n a_{ij} \log a_{ij} \quad (11)$$

is sought, subject to the constraints (1) that

$$\left. \begin{aligned} \sum_{j=1}^n a_{ij} &= \frac{R_i}{S} , \quad 1 \leq i \leq n , \\ \text{and } \sum_{i=1}^n a_{ij} &= \frac{C_j}{S} , \quad 1 \leq j \leq n \end{aligned} \right\} . \quad (12)$$

In Ref. 4, necessary conditions for an extremum are developed by Lagrange multipliers, and the unique prospective solution is $(a_{ij}) = (R_i C_j / S^2)$. Observe that $\mathcal{E}$ is a continuous function over the compact set $0 \leq a_{ij} \leq 1$ in Euclidean n^2 -space; since there is smaller contribution to the sum of Eq. (11) if an a_{ij} exceeds 1, a maximum could not be achieved at a point outside of this unit cube. Similarly, inequalities of the types $\mathcal{E}(a, b, \dots, 1, \dots) < \mathcal{E}(a, b, \dots, 1-\epsilon, \dots)$ or $\mathcal{E}(a, b, \dots, 0, \dots) < \mathcal{E}(a, b, \dots, 0+\epsilon, \dots)$, for any selected ϵ satisfying $0 < \epsilon < 1$, show that a maximum could not occur at a boundary point, either. Hence $\mathcal{E}$ has a maximum and it is, indeed, attained at an interior point of

*This form⁹ for a maximum-entropy method has been selected because $x \log x$ remains well behaved for x near zero. In fact, $x \log x \rightarrow 0^-$ as $x \rightarrow 0^+$.

the unit cube. Finally, if some R_i and/or some C_j equals zero, the corresponding terms in Eq. (11) may be omitted before the maximization process above is carried out and we therefore have

Theorem 2. Suppose $C_1, C_2, \dots, C_n, R_1, R_2, \dots, R_n$ are prescribed data which are non-negative and consistent. Let E be the solution matrix given in Eq. (6), and let M denote any other non-negative solution matrix for Eq. (4). Then

$$\mathcal{E}(M) < \mathcal{E}(E) .$$

The maximum-entropy solution E has properties that are quite different from those of the minimum-norm solution N . For example, the maximum-entropy reconstruction process is not linear, but it seems to be much less susceptible to the smearing effect. Even for test cases consisting of a core with a substantial halo component, the maximum-entropy reconstruction can yield a good core approximation. See Fig. 3 for an illustration.

Fig. 1.
A wire scanner schematic.

Fig. 2(a).
Test data generated by $7 \exp\{-x^2-y^2\} + \exp\{-(x-3)^2-(y-3)^2\}$, for $-6 \leq x, y \leq 6$ and a common mesh size of 0.3.

Fig. 2(b).
The minimum-norm reconstruction for the data shown in Fig. 2(a).

Fig. 3.
The maximum-entropy reconstruction for the data shown in Fig. 2(a).

III. SINGLE WIRE-SCAN RECONSTRUCTION--CONTINUOUS (GAUSSIAN) CORE MODEL

In Sec. II, fitting of a discrete core (i.e., a matrix) to the data collected by a single pass of a wire scanner was considered. A different approach involves approximation to these data by, instead, members of some family of continuous surfaces. In this section, a model is developed for consistent data $C_1, C_2, \dots, C_n, R_1, R_2, \dots, R_n$ and for the function family described in:

Problem 1. From among all members of the six-parameter family

$$\lambda(x,y) = a e^{\alpha(x-x_0)^2 + \beta(x-x_0)(y-y_0) + \gamma(y-y_0)^2}, \quad (13)$$

determine one which is, in some sense to be decided, an acceptable approximation to the given data. It will be shown that, with a modified, discrete least squares approximation scheme (involving the maximum-entropy solution), a unique, best surface exists and can be readily calculated.

For a well-tuned beam, a projection $D(x,y)$ in xy -space is frequently viewed as having a major component which is a bell-shaped surface (the desirable situation has a sharply focused stream of particles). Those surfaces in the family of Eq. (13) which have just such an appearance are easily characterized.

Lemma 4. Consider any function of Eq. (13) where $a > 0$, $\alpha < 0$, and $\lambda < 0$. Then, the corresponding surface $z = \lambda(x,y)$ is bell-shaped and has a unique maximum point iff $\beta^2 < 4\alpha\gamma$.

Proof. After the coordinate translation

$$\begin{aligned} X &= x - x_0 \\ Y &= y - y_0 \end{aligned} \quad (14)$$

then the rotation

$$\begin{aligned} X &= \bar{x} \cos \theta - \bar{y} \sin \theta \\ Y &= \bar{x} \sin \theta + \bar{y} \cos \theta \end{aligned}$$

through an angle determined from

$$\tan 2\theta = \frac{\beta}{\alpha - \gamma} ,$$

the form $\lambda(x,y)$ is simply

$$a e^{\bar{\alpha} \bar{x}^2 + \bar{\gamma} \bar{y}^2} . \quad (15)$$

$$\begin{aligned} \text{In Eq. (15), } \bar{\alpha} &= 1/2 \left[(\alpha + \gamma) + \sqrt{\beta^2 + (\alpha - \gamma)^2} \right] , \\ \bar{\gamma} &= 1/2 \left[(\alpha + \gamma) - \sqrt{\beta^2 + (\alpha - \gamma)^2} \right] \end{aligned} \quad (16)$$

and it is an easy matter to check that, under the stated hypotheses, $\bar{\alpha} < 0$ and $\bar{\gamma} < 0$ iff $\beta^2 < 4\alpha\gamma$. On the other hand, the function shown in Eq. (15) has a unique maximum when, and only when, it is the familiar Gaussian density, i.e., $\bar{\alpha} < 0$ and $\bar{\gamma} < 0$.

Let $\{(x_j, y_i) | 1 \leq i, j \leq n\}$ be the mesh determined in the xy -plane by one scan. For the maximum-entropy matrix $E = (R_i C_j / S)$, define

$$m = \max_{i,j} \frac{R_i C_j}{S} .$$

The set $T = \{(x_j, y_i) | R_i C_j / S = m\}$ will have k members, for some $k \geq 1$, and the numbers

$$\left. \begin{aligned} x_0^* &= \frac{1}{k} \sum_{(x_j, y_i) \in T} x_j \\ y_0^* &= \frac{1}{k} \sum_{(x_j, y_i) \in T} y_i \end{aligned} \right\} \quad (17)$$

are selected as values for two of the parameters being sought. Now, with a translation of axes and the corresponding notational changes

$$\begin{aligned} X_j &\equiv x_j - x_0^*, \quad 1 \leq j \leq n, \\ Y_i &\equiv y_i - y_0^*, \quad 1 \leq i \leq n, \end{aligned} \quad (18)$$

$$\text{and} \quad A \equiv \log a,$$

the functions shown in Eq. (13) are simply

$$\Lambda(X,Y) = \exp \{A + \alpha X^2 + \beta XY + \gamma Y^2\}.$$

Hence, the measure of approximation adopted here is

$$\sum_{i=1}^n \sum_{j=1}^n \left[\log \left\{ \frac{R_i C_j}{S} \right\} - \log \left\{ \Lambda(X_j, Y_i) \right\} \right]^2,$$

and the parameters A, α, β, γ will be determined so as to minimize

$$F(A, \alpha, \beta, \gamma) = \sum_{i=1}^n \sum_{j=1}^n \left[\log \left\{ \frac{R_i C_j}{S} \right\} - \left\{ A + \alpha X_j^2 + \beta X_j Y_i + \gamma Y_i^2 \right\} \right]^2. \quad (19)$$

Consider the linear space L of real-valued functions (matrices) $f(X,Y)$ which are defined at the n^2 grid points (X_j, Y_i) . In this space,

$$\langle f, g \rangle = \sum_{i=1}^n \sum_{j=1}^n f(X_j, Y_i) g(X_j, Y_i) \quad (20)$$

is an inner product and, consequently, the associated norm is

$$\|f\| = \sqrt{\sum_{i=1}^n \sum_{j=1}^n f^2(X_j, Y_i)}.$$

Lemma 5. Assume $n \geq 3$. The four functions

$$\phi_1(X,Y) = 1, \quad \phi_2(X,Y) = X^2, \quad \phi_3(X,Y) = XY, \quad \phi_4(X,Y) = Y^2 \quad (21)$$

are linearly independent elements in the space L .

Proof. Suppose there exist scalars a_1, a_2, a_3, a_4 such that

$$a_1 + a_2 X^2 + a_3 XY + a_4 Y^2 \equiv 0. \quad (22)$$

If, in particular, Eq. (22) is evaluated at (X_1, Y_1) , (X_1, Y_2) and (X_1, Y_3) , then pairwise subtraction among the three resulting equations shows that

$$a_3X_1 + a_4(Y_1+Y_2) = 0 ,$$

$$\text{and } a_3X_1 + a_4(Y_1+Y_3) = 0 .$$

Thus, $a_4(Y_2-Y_3) = 0$. But $Y_2 \neq Y_3$ implies that $a_4 = 0$ and hence, $a_3X_1 = 0$. If $X_1 = 0$, the above argument can be repeated with X_2 replacing X_1 so that a must be zero. A similar argument using nodes (X_j, Y_1) for $1 \leq j \leq 3$ shows that $a_1 = a_2 = 0$ as well, to complete the proof.

The linear independence of the four elements of Eq. (21) ensures existence of a minimum for the function shown in Eq. (19): for, with the element $\phi(X, Y)$ defined by $\phi(X_j, Y_i) = \log(R_i C_j / S)$, $1 \leq i, j \leq n$, the problem of finding

$$\min_{A, \alpha, \beta, \gamma} ||\phi - \{A\phi_1 + \alpha\phi_2 + \beta\phi_3 + \gamma\phi_4\}||$$

has a solution (Ref. 10, p. 137), but this precisely settles the question of minimization for Eq. (19). Since $F(A, \alpha, \beta, \gamma)$ has a minimum, it occurs at a solution of the normal equations $\frac{\partial F}{\partial A} = \frac{\partial F}{\partial \alpha} = \frac{\partial F}{\partial \beta} = \frac{\partial F}{\partial \gamma} = 0$; or, at a solution of a linear system

$$\phi P = Q \tag{23}$$

in which $P \equiv (A, \alpha, \beta, \gamma)^T$, $Q \equiv (Q_1, Q_2, Q_3, Q_4)^T$, where

$$\begin{aligned} Q_1 &\equiv n \sum_{i=1}^n \log R_i + n \sum_{j=1}^n \log C_j - n^2 \log S \\ Q_2 &\equiv \left(\sum_{j=1}^n X_j^2 \right) \left(\sum_{i=1}^n \log R_i \right) + n \sum_{j=1}^n X_j^2 \log C_j - n(\log S) \sum_{j=1}^n X_j^2 \\ Q_3 &\equiv \left(\sum_{j=1}^n X_j \right) \left(\sum_{i=1}^n Y_i \log R_i \right) + \left(\sum_{j=1}^n X_j \log C_j \right) \left(\sum_{i=1}^n Y_i \right) - \left(\sum_{j=1}^n X_j \right) \left(\sum_{i=1}^n Y_i \right) \log S \\ Q_4 &\equiv n \sum_{i=1}^n Y_i^2 \log R_i + \left(\sum_{j=1}^n \log C_j \right) \left(\sum_{i=1}^n Y_i^2 \right) - n(\log S) \sum_{i=1}^n Y_i^2 , \end{aligned} \tag{24}$$

and Φ is the matrix

$$\begin{pmatrix} n^2 & n \sum_{j=1}^n X_j^2 & \left(\sum_{j=1}^n X_j \right) \left(\sum_{i=1}^n Y_i \right) & n \sum_{i=1}^n Y_i^2 \\ n \sum_{j=1}^n X_j^2 & n \sum_{j=1}^n X_j^4 & \left(\sum_{j=1}^n X_j^3 \right) \left(\sum_{i=1}^n Y_i \right) & \left(\sum_{j=1}^n X_j^2 \right) \left(\sum_{i=1}^n Y_i^2 \right) \\ \left(\sum_{j=1}^n X_j \right) \left(\sum_{i=1}^n Y_i \right) & \left(\sum_{j=1}^n X_j^3 \right) \left(\sum_{i=1}^n Y_i \right) & \left(\sum_{j=1}^n X_j^2 \right) \left(\sum_{i=1}^n Y_i^2 \right) & \left(\sum_{j=1}^n X_j \right) \left(\sum_{i=1}^n Y_i^3 \right) \\ n \sum_{i=1}^n Y_i^2 & \left(\sum_{j=1}^n X_j^2 \right) \left(\sum_{i=1}^n Y_i^2 \right) & \left(\sum_{j=1}^n X_j \right) \left(\sum_{i=1}^n Y_i^3 \right) & n \sum_{i=1}^n Y_i^4 \end{pmatrix} \quad (25)$$

The matrix Φ has a very simple form: it is just the Gram matrix* for the four linearly independent elements $\phi_1, \phi_2, \phi_3, \phi_4$ and so, it must be¹⁰ nonsingular. Whence, Eq. (23) has a unique solution and it yields the minimum for $F(A, \alpha, \beta, \gamma)$. These results may be summarized as:

Theorem 3. Suppose $C_1, C_2, \dots, C_n, R_1, R_2, \dots, R_n$ are positive, consistent data and that $n \geq 3$. For the corresponding mesh $\{(x_j, y_i) | 1 \leq i, j \leq n\}$, let (x_0^*, y_0^*) be determined by Eq. (17) then each X_j and Y_i by Eq. (18). Form the column vector Q given in Eq. (24) and the matrix prescribed in Eq. (25), and let $(A^*, \alpha^*, \beta^*, \gamma^*)$ denote the unique solution, namely $\Phi^{-1}Q$, of Eq. (23). Then, from among all members of the family shown in Eq. (13), the one which minimizes the function in Eq. (19) is

$$\lambda(x, y) = (e^{A^*}) e^{\alpha^*(x-x_0^*)^2 + \beta^*(x-x_0^*)(y-y_0^*) + \gamma^*(y-y_0^*)^2}.$$

The scheme outlined in the preceding theorem yields a Gaussian-core, modified least squares approximation to appropriate data $C_1, C_2, \dots, C_n, R_1, R_2, \dots, R_n$. If these data were actually generated from a surface which is itself a member of Eq. (13), then it would be natural to ask when the reconstructed surface will be precisely the original one. The next lemma confirms that bell-shaped surfaces having a unique maximum and for which the correlation parameter β is zero are among those that are self-reproducing under this core reconstruction process.

*The element in the i th row and j th column is $\langle \phi_i, \phi_j \rangle$.

Lemma 6. Let $\mu(x,y) = \hat{a} \exp\{\hat{\alpha}(x-\hat{x}_0)^2 + \hat{\gamma}(y-\hat{y}_0)^2\}$ be a given surface for which $\hat{a} > 0, \hat{\alpha} < 0$ and $\hat{\gamma} < 0$, and suppose that $(\hat{x}_0, \hat{y}_0)$ is one of the nodes of the mesh $\{(x_j, y_1) | 1 \leq i, j \leq n\}$ imposed on the domain. If

$$R_i = \sum_{j=1}^n \mu(x_j, y_1) \quad , \quad 1 \leq i \leq n$$

and
$$C_j = \sum_{i=1}^n \mu(x_j, y_1) \quad , \quad 1 \leq j \leq n \quad ,$$

then the best Gaussian core approximation for the data $C_1, C_2, \dots, C_n, R_1, R_2, \dots, R_n$ will be precisely $\mu(x,y)$.

Proof. For these data, $x_0^* = \hat{x}_0$, $y_0^* = \hat{y}_0$ and

$$\frac{R_i C_j}{S} = \hat{a} e^{\hat{\alpha} X_j^2 + \hat{\gamma} Y_i^2} \quad \text{for } 1 \leq i, j \leq n \quad .$$

But then,

$$\sum_{i,j=1}^n \left[\log \left\{ \hat{a} e^{\hat{\alpha} X_j^2 + \hat{\gamma} Y_i^2} \right\} - \log \left\{ a e^{\alpha X_j^2 + \beta X_j Y_i + \gamma Y_i^2} \right\} \right]^2$$

is clearly minimized if $a = \hat{a}$, $\alpha = \hat{\alpha}$, $\beta = 0$ and $\gamma = \hat{\gamma}$, so that the core obtained is simply $\mu(x,y)$.

The Gaussian-core technique was subjected to quite extensive numerical investigation. Mainly single Gaussians and sums of two or three Gaussians were used to generate test data and then the graph of the original could be compared with the graph of its reconstruction. Data parameters varied, in addition to the number of bell-shaped components, included the coefficient(s) of the cross term(s) $(x-x_0)(y-y_0)$, the fineness of the mesh used, and the relative size of halo components. Several empirical features emerged:

- (1) If the test data were generated by a single Gaussian $\hat{a} \exp\{\hat{\alpha}(x-\hat{x}_0)^2 + \hat{\beta}(x-\hat{x}_0)(y-\hat{y}_0) + \hat{\gamma}(y-\hat{y}_0)^2\}$ then, regardless of whether $(\hat{x}_0, \hat{y}_0)$ is a node for the mesh adopted, the smaller $|\hat{\beta}|$ is, the better the approximation appeared.
- (2) For data generated by a single Gaussian, the finer the mesh, the better the approximation.
- (3) For data generated by a major Gaussian that was accompanied by one or more non-negligible halo Gaussians, the reconstruction

yielded a core that was generally "close to" the one in the data. The smaller the number and relative size(s) of halo components, the better.

Figure 4 illustrates the reconstruction for a halo that contributes some 30% of the major component.

IV. CONSECUTIVE SCANS--A MULTIVARIABLED GAUSSIAN-CORE MODEL

A. Introduction

The actual beam is a function of six or four variables according as it is bunched or not. In Secs. II and III, analytic models are introduced for an approximation to just the two-variabled viewing of the beam as it emerges at a wire scanner. Certain of the properties observed in that projection subspace lead to generalizations which permit modeling of a core for the actual beam (throughout this section, a four-variabled situation is investigated, although a bunched beam could be considered by similar techniques). Since approximations to the accelerating structure's transformations from wire scanner to wire scanner are known, the data observed by a succession of scanners may be employed collectively to determine a beam density $D(x, x', y, y')$. It is shown, in fact, that if certain generalizations of the two-variabled, maximum-entropy fits are subjected to the appropriate linear, coordinate transformations so that all the data are referred to some one particular scan position, then a modified, least squares (four-variabled) Gaussian core may be calculated there.

B. The Coordinate Transformations

Analytic descriptions of motion for charged particles are often attempted only in simplified situations, for example, when the three components of motion are assumed to be mutually independent or when the axial momentum is taken to be constant. Under such hypotheses, the real space coordinates (x, y, z) generate, in particular, separated transverse phase spaces (x, x') and (y, y') , where $x' \equiv dx/dz$ and $y' \equiv dy/dz$. Further simplification results in the modeling when descriptions of actual transport components are replaced by more elementary ones. For example, the focusing or defocusing effect of a thin quadrupole magnet, even when fringe fields and aberrations are neglected, is usually replaced by that for a drift of half the quad length, then a focusing or defocusing "instantaneous" impulse and, finally, a drift over the remaining half. Consequently, the new phase space coordinates of a point may

Fig. 4(a).

The surface $\mu(x,y) = 10e^{-(x-1.8)^2-(y-.97)^2} + e^{-x^2+xy-3y^2}$, $-2 < x, y < 2$, with a step size of 0.1.

Fig. 4(b).

The core approximation [for data generated from the surface in Fig. 4(a)]: $\gamma(x,y) = 9.53288 \exp\{-0.33837(x-1.8000)^2 + 0.02956(x-1.80000)(y-0.80000) - 1.16255(y-0.80000)^2\}$.

be obtained simply from premultiplying the old coordinate vector by a quite elementary second-order matrix whose elements approximate parameters describing the device. Such simplification is particularly convenient when points are to be studied under a succession of transformations.

Three common coordinate transformations for first-order beam dynamics on either of the two separated transverse phase spaces are:

- [i] A drift through a distance z ,

$$\mathcal{L}_z = \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix} ;$$

- [ii] A defocusing or focusing of impulse f , for a quadrupole magnet having axis length d ,

$$\mathcal{L}_{f,d} = \begin{pmatrix} 1 \pm \frac{d}{2f} & d \pm \frac{d^2}{4f} \\ \pm \frac{1}{f} & 1 \pm \frac{d}{2f} \end{pmatrix} ;$$

- and [iii] An acceleration due to an impulse Δ ,

$$\mathcal{L}_\Delta = \begin{pmatrix} 1 & 0 \\ \Delta & 1 \end{pmatrix} .$$

Thus, for example, if (x, x') phase space particles undergo first an acceleration with impulse Δ , then a focusing with impulse f and axis d , and finally a drift through the distance z , the new coordinates $(x, x')_N$ are given in terms of the old $(x, x')_0$ by

$$\begin{pmatrix} x \\ x' \end{pmatrix}_N = \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 - \frac{d}{2f} & d - \frac{d^2}{4f} \\ -\frac{1}{f} & 1 - \frac{d}{2f} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \Delta & 1 \end{pmatrix} \begin{pmatrix} x \\ x' \end{pmatrix}_0 .$$

Since these linear transformations, and hence their inverses too, are known for any segment of the linac, data collected upstream and/or downstream relative to one particular scanner may all be referred to that position for the numerical approximation procedure there.

C. Four-Variable Core Fitting

For the two-variable core obtained in Sec. III in terms of data from a single wire scan, its coefficients are determined through a least squares

fitting to the maximum-entropy reconstruction generated by those data. A somewhat similar approach is taken here: An intermediate four-variable Gaussian is constructed for each wire scanner by a modified, least squares fit to its generalized maximum-entropy model, then neighbors of a selected scanner have their intermediate Gaussians transformed (via the known coordinate transformations) to the selected position so that a permanent Gaussian core may be fitted there. Each Gaussian fitting requires the solution of a nonsingular algebraic system in eleven unknowns.

Consider, then, the projection of a beam observed as the consistent data $C_1, C_2, \dots, C_n, R_1, R_2, \dots, R_n$ at one scan position. The corresponding partition $\{x_1 < x_2 < \dots < x_n, y_1 < y_2 < \dots < y_n\}$ of xy subspace can be extended to one for (x, x', y, y') space, say

$$\{(x_j, x'_k, y_i, y'_\ell) \mid 1 \leq i, j, k, \ell \leq n\} \quad (26)$$

by selecting x'_1 and x'_n to be a theoretical* minimum and maximum for dx/dz then partitioning the x' -axis by points $x'_k = x'_1 + (x'_n - x'_1)(k-1)/(n-1)$ for $1 \leq k \leq n$ (with a similar scheme for y'). Subsequently, there are many possible extensions of the xy maximum-entropy matrix considered in Sec. III; a simple one is

$$\hat{E} = (\hat{E}_{jki\ell}) \triangleq \left[\frac{R_i C_j}{n^2 S} \right], \quad 1 \leq i, j, k, \ell \leq n, \quad (27)$$

and its discrete projection surely conforms in xy subspace since

$$\sum_{k=1}^n \sum_{\ell=1}^n \hat{E}_{jki\ell}$$

is precisely Eq. (6). The intermediate analytic model for Eq. (27) will be developed first.

Let $\phi_i(x, z, y, w)$, $1 \leq i \leq 11$, be the eleven functions defined on Euclidean 4-space by:

$$\begin{aligned} \phi_1 &\equiv 1, \quad \phi_2 \equiv x^2, \quad \phi_3 \equiv xy, \quad \phi_4 \equiv y^2, \quad \phi_5 \equiv xz, \\ \phi_6 &\equiv z^2, \quad \phi_7 \equiv xw, \quad \phi_8 \equiv w^2, \quad \phi_9 \equiv yz, \quad \phi_{10} \equiv yw, \\ \phi_{11} &\equiv zw, \end{aligned} \quad (28)$$

*One such choice is $x'_1 = -6.0$ milliradians and $x'_n = 6.0$ milliradians.

and consider the problem of determining parameters $a, x_0, x'_0, y_0, y'_0, \delta_2, \delta_3, \dots, \delta_{11}$ so as to fit

$$a \exp \left[\sum_{r=2}^{11} \delta_r \phi_r(x-x_0, x'-x'_0, y-y_0, y'-y'_0) \right]$$

to the data represented by Eq. (27) in some way. Select

$$x'_0{}^* = \frac{x'_1 + x'_n}{2}, \quad y'_0{}^* = \frac{y'_1 + y'_n}{2} \quad (29)$$

and x_0^* and y_0^* as in Eq. (17). Now with a translation of coordinate axes and the corresponding notational changes

$$X_j = x_j - x_0^*, \quad Y_i = y_i - y_0^*, \quad X'_k = x'_k - x'_0{}^*, \quad Y'_\ell = y'_\ell - y'_0{}^*, \quad (30)$$

and

$$A = \log a,$$

development proceeds quite analogously to that for Theorem 3, but with the inner product

$$\langle f, g \rangle = \sum_{i,j,k,\ell=1}^n f(X_j, X'_k, Y_i, Y'_\ell) g(X_j, X'_k, Y_i, Y'_\ell) \quad (31)$$

and corresponding norm $\|f\| = \sqrt{\langle f, f \rangle}$.

(a) If $n \geq 3$, the eleven functions shown in Eq. (28) are linearly independent elements in the space having inner product expressed by Eq. (31).

(b) With $\phi(X, X', Y, Y')$ defined by $\phi(X_j, X'_k, Y_i, Y'_\ell) = \log \left(\hat{E}_{jki\ell} \right)$,

the function

$$\begin{aligned} F(a, \delta_2, \dots, \delta_{11}) &\stackrel{d}{=} \sum_{i,j,k,\ell=1}^n \left[\log (\hat{E}_{jki\ell}) - \log \left\{ a \exp \left(\sum_{r=2}^{11} \delta_r \phi_r(X_j, X'_k, Y_i, Y'_\ell) \right) \right\} \right]^2 \\ &= \left\| \phi - \left\{ A\phi_1 + \sum_{r=2}^{11} \delta_r \phi_r \right\} \right\|^2 \end{aligned}$$

has a minimum.

(c) The eleven normal equations $\partial F / \partial A = 0$ and $\partial F / \partial \delta_s = 0$ for $s = 2, 3, \dots, 11$ again take the same simple form: the i th equation is just

$$\langle \phi_1, \phi_1 \rangle A + \langle \phi_1, \phi_2 \rangle \delta_2 + \dots + \langle \phi_1, \phi_{11} \rangle \delta_{11} = \langle \phi_1, \phi \rangle, \quad 1 \leq i \leq 11. \quad (32)$$

(d) The coefficient matrix in the linear system shown in Eq. (32) is nonsingular.

Theorem 4. Suppose $C_1, C_2, \dots, C_n, R_1, R_2, \dots, R_r$ are positive, consistent data at one wire scan and $n \geq 3$. For the corresponding mesh [Eq. (26)], determine $x_o^*, y_o^*, x_o'^*, y_o'^*$ according to Eq. (29) and then X_j, X_k', Y_i, Y_ℓ' as in Eq. (30). Let $(A^*, \delta_2^*, \dots, \delta_{11}^*)^T$ be the solution of the linear system [Eq. (32)]. Then the best, modified, least squares, intermediate Gaussian-core approximation at this wire scanner is

$$G \left[\begin{pmatrix} x \\ x' \\ y \\ y' \end{pmatrix} \right] = (e^{A^*}) \exp \left\{ \sum_{r=2}^{11} \delta_r^* \phi_r (x-x_o^*, x'-x_o'^*, y-y_o^*, y'-y_o'^*) \right\} \quad (33)$$

For a chosen wire scanner, say WS#0, we seek a Gaussian core $D(x, x', y, y')$ there, but one which also incorporates the observed data at some prescribed neighboring scanners WS#1, WS#2, ..., WS#K. For each $m = 1, 2, \dots, K$, the coordinate transformation, $\mathcal{L}^{(m)}$, referring the position for WS#m to that of WS#0 is known: for instance, if WS#m were at a position obtained simply by one drift, in the amount z , downstream from position #0, then

$$\begin{pmatrix} x \\ x' \\ y \\ y'_m \end{pmatrix} = \begin{pmatrix} 1 & z & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ x' \\ y \\ y' \end{pmatrix}_0,$$

so that, here,

$$\mathcal{L}^{(m)} = \begin{pmatrix} 1 & z & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & z \\ 0 & 0 & 0 & 1 \end{pmatrix}^{-1}.$$

Hence, in this illustration the intermediate Gaussian core $G_m[(x, x', y, y')^T]$ at WS#m would be redescribed at WS#0 simply as

$$G_m[\mathcal{L}^{(m)}(x, x', y, y')^T] = G_m[(x-zx', x', y-zy', y')^T].$$

(In general, an $\mathcal{L}^{(m)}$ is a product of, or inverse of a product of, the listed coordinate transformations according as WS#m is upstream, or downstream, from WS#0.)

Now to start, the four parameters $\hat{x}_0, \hat{x}'_0, \hat{y}_0, \hat{y}'_0$ in the trial family

$$D(x, x', y, y') = b \exp \left\{ \sum_{r=2}^{11} \mu_r \phi_r(x - \hat{x}_0, x' - \hat{x}'_0, y - \hat{y}_0, y' - \hat{y}'_0) \right\} \quad (34)$$

are easy to fix since they may be selected as the $x_0^*, x_0'^*, y_0^*$ and $y_0'^*$ one calculates for the intermediate Gaussian approximation at WS#0. Then, after axes translation

$$X = x - x_0^*, X' = x' - x_0'^*, Y = y - y_0^*, Y' = y' - y_0'^*$$

and the notational changes

$$X_j = x_j - x_0^*, X'_k = x'_k - x_0'^*, Y_l = y_l - y_0^*, Y'_\ell = y'_\ell - y_0'^*, \quad (35)$$

and $B = \log b$,

D can be determined as that member of Eq. (34) which best fits the average of the transforms of the neighboring intermediate cores. In other words,

$$\text{if } G^A(x, x', y, y') \equiv \frac{1}{K} \sum_{m=1}^K G_m \left[\mathcal{L}^{(m)}(x, x', y, y')^T \right], \quad (36)$$

the best $D(x, x', y, y')$ from family [Eq. (34)] is the one for which

$$\begin{aligned} \mathcal{F}(b, \mu_2, \dots, \mu_{11}) &\stackrel{d}{=} \sum_{i,j,k,\ell=1}^n \left[\log \left(G^A(x_j, x'_k, y_i, y'_\ell) \right) - \log \left\{ D(x_j, x'_k, y_i, y'_\ell) \right\} \right]^2 \\ &= \left\| \psi - \left\{ B\phi_1 + \sum_{r=2}^{11} \mu_r \phi_r \right\} \right\|^2 \end{aligned}$$

is minimized*; here, of course, $\psi(X, X', Y, Y')$ denotes the inner product space element given by

$$\psi(X_j, X'_k, Y_i, Y'_\ell) = \log \left\{ G^A(x_j, x'_k, y_i, y'_\ell) \right\}. \quad (37)$$

This final portion of the development is summarized as:

Theorem 5. For each $m = 1, 2, \dots, K$, let $G_m(x, x', y, y')$ be the intermediate Gaussian-core function [Eq. (33)] corresponding to a prescribed wire scanner WS#m with observed, positive, consistent data $C_1^{(m)}, C_2^{(m)}, \dots, C_n^{(m)}, R_1^{(m)}, R_2^{(m)}, \dots, R_n^{(m)}$.

Assume that $n > 3$. Let WS#0 be a fixed scanner whose intermediate Gaussian core

*Once again, linear independence of the functions (28) ensures existence and uniqueness of a solution to this problem.

has center $(\hat{x}_0, \hat{x}'_0, \hat{y}_0, \hat{y}'_0)$, and let $\mathcal{L}^{(m)}$ ($1 \leq m \leq K$) denote the coordinate transformation relating the position of WS#m to that of WS#0. For the partition [Eq. (26)] determined at WS#0 and the corresponding Eq. (35), let $(B^*, \mu_2^*, \dots, \mu_{11}^*)^T$ be the solution of the linear system

$$\langle \phi_1, \phi_1 \rangle B + \langle \phi_1, \phi_2 \rangle \mu_2 + \dots + \langle \phi_1, \phi_{11} \rangle \mu_{11} = \langle \phi_1, \psi \rangle, \quad 1 \leq i \leq 11,$$

where ψ is given by Eqs. (36) and (37). Then a Gaussian beam core at WS#0 which incorporates the data from WS#1, WS#2, ..., and WS#K is the function

$$D(x, x', y, y') = (e^{B^*}) \exp \left\{ \sum_{r=2}^{11} \mu_r^* \phi_r(x - \hat{x}_0, x' - \hat{x}'_0, y - \hat{y}_0, y' - \hat{y}'_0) \right\}. \quad (38)$$

It is clear that the core Eq. (38) is only one of many possible; for instance, minimization of, say,

$$\sum_{m=1}^K || \log \left(G_m[\mathcal{L}^{(m)}] \right) - \log\{D\} ||^2$$

might be considered instead, or alternative generalizations for the maximum-entropy matrix could be used in place of $\hat{E}$ for the first stage. These and other related questions are currently being investigated.

ACKNOWLEDGMENTS

The author is pleased to acknowledge the many informative and stimulating discussions with members of, and visitors to, Group MP-9. I particularly wish to thank R. A. Jameson for his patient guidance and assistance during my term as a visiting staff member at LAMPF. It has been a most rewarding experience.

REFERENCES

1. R. W. Goodwin, E. R. Gray, G. M. Lee, and M. F. Shea, "Beam Diagnostics for the NAL 200-MeV Linac," in Proc. 1970 Proton Linear Accelerator Conf., National Accelerator Laboratory (October 1970).
2. T. F. Budinger and G. T. Gullberg, "Three-dimensional Reconstruction in Nuclear Medicine by Iterative Least-Squares and Fourier Transform Techniques," Lawrence Berkeley Laboratory report LBL-2146 (January 1974).
3. G. T. Herman and A. Lent, "Quadratic Optimization for Image Reconstruction, Part I," State University of New York at Buffalo, report 103 (August 1975).

4. G. N. Minerbo and J. G. Sanderson, "Reconstruction of a Source from a Few (2 or 3) Projections," Los Alamos Scientific Laboratory report LA-6747-MS (March 1977).
5. H. J. Trussell, "Improved Methods of Maximum A Posteriori Image Restoration," Los Alamos Scientific Laboratory report LA-6516-T (September 1976).
6. H. Koziol, "Evaluation of Near-Gaussian Distributions with Ill-defined Tails," Los Alamos Scientific Laboratory, internal communication.
7. K. L. Manders, "Linear and Non-linear Computation: An Analysis of Linac Beam Trajectory Simulation Programs," European Organization for Nuclear Research report CERN/MPS/LIN/72-3 (August 1972).
8. A. Albert, Regression and the Moore-Penrose Pseudoinverse (Academic Press, New York, 1972).
9. L. R. D'Addario, "Maximum A Posteriori Probability and Maximum Entropy Reconstruction, in Image Processing for 2-d and 3-d Reconstruction from Projections," Stanford University and the Optical Society of America, Technical Digest (August 1975).
10. P. J. Davis, Interpolation and Approximation (Blaisdell Publishing Co., Inc., Lexington, Massachusetts, 1963).
11. B. Bru and D. J. Warner, "Bunch 74: An Improved Multiparticle Beam Simulation Program," European Organization for Nuclear Research report CERN/MPS/LIN/75-2 (December 1975).