JOURNAL OF COMPUTATIONAL PHYSICS 28, 408–415 (1978)

Calculations Involving Ion Beam Source*

J. C. WHITSON AND J. SMITH

Computer Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830

AND

J. H. WHEALTON

Fusion Energy Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830

Received May 17, 1977; revised November 15, 1977

A convergent numerical model for the calculation of the optical properties of an ion beam is given. The ion beam is formed by extracting ions from a plasma and subsequently accelerating these ions with an electrode system. The model includes the effects of space-charge of the ions and an equilibrium distribution of electrons. Methods are given for determining the existence of the solution to the nonlinear difference equations, and a convergent iterative numerical procedure is described. Comparisons are made with a procedure that previously has been used to solve such a model.

1. INTRODUCTION

Neutral beam injection systems for heating and sustaining controlled thermonuclear reactor (CTR) plasmas require the production of hydrogen and deuterium ion beams of high current density and minimum beam divergence. These beams are prepared by extracting ions from a weakly ionized gas discharge and accelerating these ions with an electrode system [1, 2]. The space-charge of the ion beam is neutralized immediately as it exits from the accelerator by charge exchange in a neutralizer filled with gas. We consider a single beamlet in such an ion source. Figure 1 shows a cross section of the region where we model the electric fields. We discuss a correct method for solving the model equations in the extraction region of such an ion source.

In Section 2 we discuss the mathematical model for the beamlet in the electrode region of the ion source. Section 3 details a mathematical proof of the existence and uniqueness of the numerical solution to the model. In Section 4 we compare the results of the present method of solution to a previously used method.

* Research sponsored by the Energy Research and Development Administration under contract with Union Carbide Corporation. The U.S. Government's right to retain a nonexclusive royalty-free liscense in and to the copyright covering this paper, for governmental purposes, is acknowledged.

0021-9991/78/0283-0408$2.00/0
Copyright © 1978 by Academic Press, Inc.
All rights of reproduction in any form reserved.

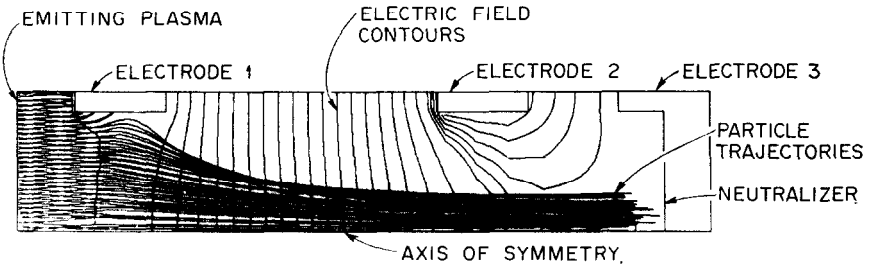

FIG. 1 A solution obtained by method II.

2. THE MODEL

The model requires the simultaneous solution of two time-independent partial differential equations. Poisson's equation must be solved for the electrostatic potential $\phi(\mathbf{r})$ in the region of interest and Vlasov's equation must be solved for the ion distribution $f_1(v)$, or alternatively the ion density $n_i = \int f_i(v)\, dv$, in the same region. A Boltzmann distribution of electron densities is assumed to exist throughout the region. Specifically, we consider the solution of

$$\nabla^2\phi = -(e/\epsilon_0)\{n_i - n_{e0} \exp[-e(\phi_0 - \phi)/kT_e]\}, \tag{1}$$

where ϕ is the electric potential, ϵ_0 is the permittivity of free space, e is the electron charge, k is Boltzmann's constant, ϕ_0 is the value of the potential at the center of the plasma, n_i is the ion density, and n_{e0} is the electron density at the center of the plasma. The Vlasov equation is solved indirectly by computing ion trajectories in the potential ϕ. Specifically, n_i is computed by the following procedure.

A nonrelativistic Hamiltonian for a particle with mass m_i, charge eZ, and momentum $\mathbf{P}^2/2m_i + eZ\phi$. Hamilton's equations of motion of the particle are then given by $\partial H/\partial q_i = -dp_i/dt$ and $\partial H/\partial p_i = dq_i/dt$, where P_i represents the generalized momentum and q_i its corresponding generalized coordinate. We numerically integrate Hamilton's equations using an ordinary differential equation integrator with given initial conditions. This procedure gives a path that ions starting from a given point with a given initial momenta would follow. Once paths have been calculated for a set of initial conditions, we calculate the effective charge in a cell of the mesh where Eq. (1) has been differenced. The charge is then deposited on the cell verticles. Letting I represent the fraction of the total current in the beamlet carried by a single path, the effective charge in a given cell through which the trajectory passes can be written as $q = \int_c \mathrm{I}\, d\mathbf{l}/|\, V\,|$, where the integration is along the particles path through the cell. In other words, $Q = I\Delta t$, where Δt is the time spent in each cell by the particle.

The initial conditions for the equations of motion are chosen by assuming the ion distribution in the source plasma to be a drifting Maxwellian with a drift velocity toward the accelerating electrode region. An average initial velocity is computed for

each of the paths to be calculated using this distribution. The starting position is chosen arbitrarily by assigning an initial coordinate in the source plasma. In practice we have found that the results of the code are not sensitive to the initial conditions of the trajectories nor the exact procedure of depositing the charge on the cell vertices. Our procedure for analyzing an electrode system is as follows. First, Eq. (1) is solved with $n_i = 0$. A space-charge n_i is then obtained by tracing ion trajectories and Eq. (1) is solved with this new n_i. The process is continued until no significant change in the trajectories is occurring. This process has been used before by other investigators [1–3]. The numerical solution of Eq. (1) with a given n_i and specified boundary conditions has proved to be a troublesome task. In Section 3, we show that the approximate solution to this nonlinear equation exists for any n_i, is unique, and can be obtained from any starting value of ϕ. While this settles the convergence question for Eq. (1) with fixed n_i, the question of convergence of the process of updating n_i cannot be answered at this point. All of our numerical experiments indicate that this process does converge.

We choose to write the coordinate lengths in dimensionless variables by letting

$$u = e(\phi_0 - \phi)/kT_e\,, \tag{2}$$

$$\chi_i = x_i/\lambda_D\,, \tag{3}$$

$$\lambda_D = [\epsilon_0 kT_e/(n_{e0}e^2)]^{1/2}, \tag{4}$$

where λ_D is the electron Debye length at the center of the source plasma. We rewrite Eq. (1) as

$$\nabla^2 u = \rho_i - e^{-u}, \tag{5}$$

where

$$\rho_i = n_i/n_{e0}\,.$$

The boundary condition $n_{e0} = n_i$ at the plasma source boundary is used to ensure space-charge neutrality. The potential u or $\partial u/\partial x$ is specified on all boundaries.

Equation (5) may be approximated by a set of suitable finite-difference equations. For simplicity, we assume that we have a two-dimensional grid system. Using standard central difference approximations for the partial derivatives, one has an equation for each mesh point of the form

$$u_0 = c_1u_1 + c_2u_2 + c_3u_3 + c_4u_4 + c_5\,, \tag{7}$$

where u_0 is the value of the potential at the mesh point under consideration. u_1, u_2, u_3 and u_4 are the values of the potential at the neighbors of u_0; c_1, c_2, c_3, and c_4 are strictly geometry-dependent; and c_5 depends on the right side of Eq. (5) and the geometry of the problem. The system of equations arising from Eq. (7) can be written in the form (see Ref. [4])

$$Au = b, \tag{8}$$

where A is an $n \times n$ matrix, u is a column matrix with the values of u at each mesh point as elements, and b is an $n \times 1$ matrix whose elements are formed by c_5. Two methods of solving the system of equations are compared. In Method I, described in detail elsewhere [5], an attempt is made to solve the system of equations with the assumption that c_5 is a constant. Then the system of linear equations is solved, a new c_5 is calculated, and the equations are solved again. In practice, c_5 depends on under-relaxation parameters that control the amount of change in c_5 from one update to another. One can think of Method I as follows. We solve or partially solve Eq. (8) for u; then using this value of u we calculate a new b and repeat the process until the full nonlinear system is satisfied by the last value of u. The solution is obtained in practice by using Eq. (7) as the basis for a relaxation scheme.

In Method II, which is the subject of this paper, we solve the set of equations in Eq. (7) in their full nonlinear form. That is, we solve for u_0 in the equation

$$u_0 - Ge^{-u_0} = c_1u_1 + c_2u_2 + c_3u_3 + c_4u_4 + c_5', \tag{9}$$

where c_5' is now composed of ρ_i and G is a geometrical factor. We note that c_5' is now a constant and does not depend on u in our scheme. Because it is derived from the ion space-charge, it is calculated from the previous u. Thus the relaxation scheme is now based on Eq. (9).

This system of equations can be written as

$$Au + \Phi(u) = 0, \tag{10}$$

where again A is an $n \times n$ matrix, u is a column matrix, and $\Phi(u)$ is a column matrix whose elements are formed from the nonlinear term and c_5'. We show in the next section the conditions under which the system of Eq. (10) has a unique solution and how to obtain that solution.

3. Method of Solution

In this section we show how one can solve a set of n nonlinear equations in n unknowns. The restrictions on the nonlinearity will become obvious as we proceed. The system of equations is given by

$$Au + \Phi(u) = 0, \tag{11}$$

where A is an $n \times n$ symmetric matrix and $\Phi(u)$ is an $(n \times 1)$ column matrix of nonlinear elements. We define a function F as

$$F(u_1, u_2, \ldots, u_n) = \tfrac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} u_i a_{ij} u_j + \sum_{i=1}^{N} \Psi_i(u_i), \tag{12}$$

where a_{ij} is a typical element of A and

$$\Psi_i(u_i) = \int_{\alpha}^{u_i} \phi_i(t)\, dt. \tag{13}$$

The term $\phi_i(t)$ is a typical element of the nonlinear column vector $\Phi(u)$, and α is arbitrary.

First, we note that a solution to the system of equations given by Eq. (11) can be obtained by finding the critical values of F:

$$\frac{\partial F}{\partial u_n} = \sum_{j=1}^{N} a_{nj}u_j + \phi_n(u_n) = 0. \tag{14}$$

One can establish the nature of the critical points of F by calculating the Hessian matrix defined by

$$H_{nm} = \partial^2 F/\partial u_n \partial u_m\,. \tag{15}$$

If the Hessian is positive definite in a given region, the function F has a local minimum in that region [6]. If H is uniformly positive definite, then F has only one minimum and is uniformly convex [6]. For the function defined by Eq. (12)

$$H_{nm} = a_{nm} + \phi_n'(u_n)\, \delta_{nm}\,, \tag{16}$$

where $\phi_n'(u_n)$ is the first derivative of ϕ, and δ_{nm} is the Kronecker delta. Thus, the question of whether we have a unique solution to Eq. (11) can be answered by examination of the Hessian matrix given by Eq. (16).

Assuming that one has established that H is uniformly positive definite, and this will be the case if A is positive definite and $\phi_n'(u_n) \geqslant 0$, then a unique solution to Eq. (11) exists. A prescription for finding that solution is now given. Let u_i^k represent an approximation for u_i and u_i^{k+1} represent a successive approximation to u. We need to establish that

$$F(u_1\,, u_2\,,..., u_i^k,..., u_n) - F(u_1\,, u_2\,,..., u_i^{k+1},..., u_n) > 0, \tag{17}$$

or we need a prescription for successively estimating u_i such that we always move toward the minimum of F. Let

$$\frac{\partial F}{\partial u_i} = f_i(u_1{}^k, u_2{}^k,..., \hat{u}_i{}^k,..., u_n{}^k) = 0, \tag{18}$$

$$\hat{u}_i{}^k = g_i(u^k,..., u_{i-1}^k\,, u_{i+1}^k\,,..., u_i{}^k), \quad \text{and} \tag{19}$$

and

$$u_i^{k+1} = u_i{}^k + \omega(\hat{u}_i{}^k - u_i{}^k), \tag{20}$$

where ω is presently an unknown constant parameter and $(\hat{u}_i{}^k - u_i{}^k)$ will be called

the residual denoted by r_l. The u_l^{k+1} approximation to u_l is obtained by solving Eq. (18), represented by Eq. (19), and then using Eq. (20). We have

$$F(u_1,\ldots,u_l^k,\ldots,u_n) - F(u_1,\ldots,u_l^{k+1},\ldots,u_n) = -\int_{u_l^k}^{u_l^k+\omega r_l} f_l(u_1,\ldots,u_{l-1},x,u_{l+1},\ldots,u_n)\,dx. \tag{21}$$

Because

$$f_l(u_1,\ldots,u_n) = \sum_{j=1}^{N} a_{lj}u_j + \phi_l(u_l) = a_{ll}u_l + \sum_{j\neq l}^{N} a_{lj}u_j + \phi_l(u_l), \tag{22}$$

we can substitute Eq. (22) into Eq. (21), using Eq. (19), and integrate. We obtain

$$F(u_1,\ldots,u_l^k,\ldots,u_n) - F(u_1,\ldots,u_l^{k+1},\ldots,u_n) = \frac{a_{ll}}{2}(\omega)(2-\omega)\,r_l^2 + \omega r_l\phi_l(u_l^k + r_l) - \int_{u_l^k}^{u_l^k+\omega r_l}\phi_l(t)\,dt. \tag{23}$$

The question of convergence of the scheme defined by Eqs. (19) and (20) now becomes one of examination of the right side of Eq. (23). If the right side of Eq. (23) is positive, the scheme converges. If a_{ll} is positive, the first term is positive for any $0 < \omega < 2$. We need only to consider

$$E = \omega r_l\phi_l(u_l^k + r_l) - \int_{u_l^k}^{u_l^k+\omega r_l}\phi_l(t)\,dt. \tag{24}$$

One can write Eq. (24) as

$$E = \int_{u_l^k}^{u_l^k+\omega r_l}[\phi_l(u_l^k + r_l) - \phi_l(t)]\,dt. \tag{25}$$

If $\omega \leqslant 1$ and $\phi_l'(t) \geqslant 0$, then for any finite value of r_l one can see from Eq. (25) that $E \geqslant 0$. Thus, one has global convergence of the scheme for $\omega \leqslant 1$. The above argument can be presented in formal mathematical language; the interested reader should see Ref. [6]. We note that there are differences between the above argument and the proof found in Ref. [6]. For instance, the matrix A above, although symmetric, does not have to be an M matrix, and one can have global convergence with an $\omega > 1$ for some problems. For a definition of the M matrix and its importance in matrix analysis, the reader is referred to Ref. [3]. By using Eq. (24), examination of a specific case can be made. We have investigated the case $\phi(u) = u^3$ and find that ω can be greater than 1. For the partial differential equation $-\nabla^2 u - e^{-u} + \rho_i = 0$, we have the matrix A formed by standard approximations for $-\nabla^2 u$ and $\phi_i(u) = -e^{-u} + \rho_i$. In Ref. [4] it is shown that A is real and symmetric, with positive diagonal entries and nonpositive off-diagonal entries, and A is positive definite. In our iteration scheme, ρ_i is a constant term, being derived from the previous u; thus $\phi_i'(u) \geqslant 0$ for all u.

The above results assure us that a solution to the nonlinear difference equations for our problems exists and is unique for any ρ_i. Moreover, we have an efficient iterative scheme to find that solution.

In the actual computer code, we use Newton's method to solve each nonlinear Eq. (18) as indicated by Eq. (19). We start the process with $\omega = 1$ for a few iterations until r_i becomes sufficiently small and then increase ω to approximately 1.7. This process is consistent with the above results and in practice has been found to speed the convergence of the scheme. In Section 4 we show some typical comparisons of computational work versus residual size.

4. Results of the Two Methods

In Fig. 2, we have plotted computational work against the absolute residual for the two methods. The abscissa for Method I is the total number of point successive overrelaxation (SOR) passes; for Method II the abscissa is the number of SOR passes plus the number of Newton iterations divided by the number of mesh points. The

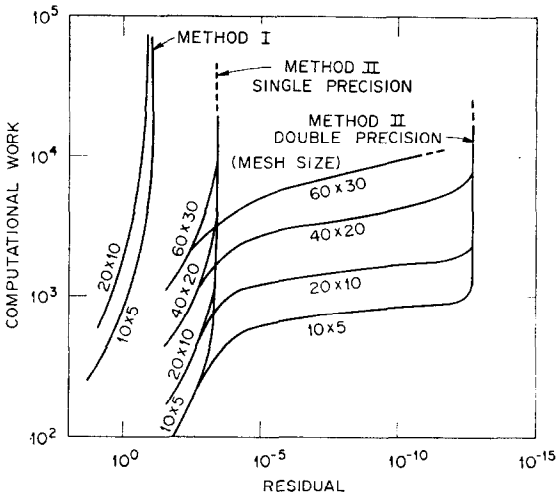

Figure 2

ordinate in each case is the maximum absolute residual. The curve for Method I was calculated in IBM 360 single precision (32-bit work length). Curves for Method II were calculated in both single- and double-precision arithmetic. The number $[n \times n]$ labeling each curve is the mesh size for that part of the curve. One can readily see that Method I does not do a very good job with this problem. While Method II is successful, double-precision arithmetic is necessary on the IBM machines to establish without doubt the rapid convergence of the Newton-SOR scheme. It seems noteworthy to mention that we first executed this code on a CDC 7600 computer. The code was fast and stable in single precision on that machine, because of its long word

length and an exponential function that is calculated with a fast table look-up. The speed of the exponential evaluation is critical in this scheme because one must solve each nonlinear equation at every mesh point by use of Newton's method. It was necessary to write a table look-up and use double-precision arithmetic for the SOR iteration in order to duplicate the CDC results on an IBM machine.[1]

ACKNOWLEDGMENT

The authors would like to thank E. F. Jaeger and L. D. Stewart for many helpful discussions concerning this project.

REFERENCES

1. K. HALBACK, "Mathematical Models and Algorithms for the Computer Program 'Wolf,'" LBL-4444, Berkeley, Calif., 1975.
2. E. F. JAEGER AND J. C. WHITSON, "Numerical Simulation for Axially Symmetric Beamlets in the DuoPIGatron Ion Source," ORNL/TM-4990, Oak Ridge, Tenn., 1975.
3. D. G. BATES, "A Computer Program for the Design of High Perveance Ion Sources," CLM-R53, Abingdon, Oxfordshire, England, 1966.
4. R. S. WARGA, "Matrix Iterative Analysis," Prentice–Hall, Englewood Cliffs, N. J., 1962.
5. J. H. WHEALTON, E. F. JAEGER, AND J. C. WHITSON, *J. Computational Physics* **27** (1978), 32–41.
6. J. M. ORTEGA AND W. D. RHEINBOLDT, "Iterative Solution of Nonlinear Equations in Several Variables," Academic Press, New York, 1970.
7. L. BERS, *J. Res. Nat. Bur. Stand. Sect. B M. M. Ph.* **51** (1953), 229–236.
8. W. KAPLAN, "Advanced Calculus," 2nd ed., Addison–Wesley, Reading, Mass., 1973.

[1] For further information the reader is referred to Refs. [7, 8].