JOURNAL OF COMPUTATIONAL PHYSICS 76, 201–230 (1988)

Back-Averaging: An Accelerated Iterative Method for Simulating Plasma Diffusion

MICHAEL E. KRESS*

Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, New York 10012

Received December 30, 1986, revised April 30, 1987

A fast accurate multi-stage numerical method, back-averaging, is developed, analyzed, and optimized The technique is used to solve an alternating dimensional numerical simulation of fusion plasma transport based on nonlinear resistive magnetohydrodynamics (MHD) equations The geometry of the model is a complicated time-dependent two-dimensional configuration of flux contours analogous to a doublet Numerical convergence rate comparisons of the optimized back-averaging method with various other iterative techniques for solving the numerical problem show that optimized back-averaging is the fastest method of all those considered Moreover, to accomplish convergence in a practical length of time for extremely peaked profiles and complicated time varying configurations, back-averaging is essential Further, with a minimum of additional computation, optimized back-averaging yields extreme accuracy

INTRODUCTION

This is a theoretical and numerical study which develops a practical algorithm for solving nonlinear resistive MHD equations used in transport theory of fusion plasmas. The method is applicable to systems of equations other than the MHD system treated here.

The derivation of the system we treat, Grad and Hogan [1], is based on the observation that a resistive, heat conducting MHD fluid evolves on several different time scales. Considerable advantage can be gained if physical phenomena which evolve on a fast time scale can be held fixed while those evolving on a slow time scale vary. These considerations were first applied in Grad and Hogan [1] diffusion theory. There, it was recognized that the time derivative in the momentum equation serves only to facilitate the distribution of plasma pressure over the flux contours of the magnetic field. We omit this time derivative from the governing system of equations and lose plasma waves but obtain a vastly simpler set of equations which are interpreted as follows: the magnetic field evolves on a time scale governed by the resistivity and the plasma pressure is instantaneously adjusted over the flux

* Permanent address College of Staten Island, City University of New York, 130 Stuyvesant Place, Staten Island, New York 10301

0021-9991/88 $3 00

contours which control the magnetic field. Thus, pressure balance occurs on a time scale which is not included in the final system of equations.

More recent approaches to the calculation of transport in toroidal plasma fusion devices include variations of the original Grad–Hogan theory. In particular, the papers of Pao [2] and Jardin [3] are notable since, for certain systems, they provide a straightforward route toward numerical implementation. However, for plasma systems which contain complex three-dimensional magnetic field topology, Grad's flux surface averaged technique of preparing the governing equations, provides an optimal mix of physics and geometry. Here we use Grad's flux surface averaged equations and develop a fast, accurate iterative method for numerical solution. A detailed treatment of the equations and the solution algorithm (Alternating Dimension Algorithm) appears in [4–7].

The introduction of flux surface averaging leads to a novel generalised or "queer" differential equation [8, 9]. Numerical solutions of the flux surface averaged equations have been carried out over the past decade by standard iteration techniques. These techniques fail for the two-dimensional configurations, involving complicated time varying topology including singularities, which we treat here. For these, we establish and accelerate convergence by multi-stage iterative techniques, in particular, back-averaging. The technique applies generally to nonlinear iterative schemes for finding a fixed point.

In Sections 1, 2, and 3, we outline the Alternating Dimensional (AD) algorithm for a resistive plasma and describe its numerical formulation for a particular magnetic field topology. Then, we define back-averaging formally and discuss regions of convergence and optimization. Finally, we demonstrate the advantages of back-averaging by using it in the Alternating Dimension numerical code and compare the calculated convergence rates with those of other methods.

1. The System of Equations

We solve the system of equations:

$$\Delta\psi = J(\psi, t), \tag{1}$$

$$\frac{\partial\psi}{\partial t} + \bar{U}\cdot\nabla\psi = \eta\,\Delta\psi, \tag{2}$$

$$\nabla\cdot\bar{U} = 0. \tag{3}$$

in two dimensions (x, y) and time t (see Fig. 1). This system describes the evolution of a resistive plasma, which passes at each instant through an equilibrium state for a large aspect ratio Tokamak in the range of low β (the ratio of plasma pressure to magnetic pressure). Here η is a constant resistivity, $\psi(x, y, t)$ is the magnetic flux, $J(\psi, t)$ is transverse current density, and $\bar{U}(x, y, t)$ is plasma velocity [6].

Following [5], we introduce the microcanonical area-weighted average on a flux surface,

$$\langle \phi \rangle = \oint \frac{\phi \, ds}{|\nabla V|}, \tag{4}$$

where the line integral is taken along a $\psi =$ constant contour, and $V(\psi)$ is the area enclosed by the contour.

We define the inductance $K(V)$ [5, 10] as

$$K(V) = \langle |\nabla V|^2 \rangle. \tag{5}$$

Taking the flux average of Eq. (2) (recall that η is constant), we obtain

$$\left\langle \frac{\partial \psi}{\partial t} \right\rangle + \langle \bar{U} \cdot \nabla \psi \rangle = \eta \langle \Delta \psi \rangle. \tag{6}$$

An elementary calculation shows that

$$\langle \Delta \psi \rangle = (K \psi')', \tag{7}$$

where ψ' indicates the derivative of $\psi(V)$ with respect to V, ψ being the inverse function associated with $V(\psi)$ [5].

From Eqs. (6), (7), and (3)

$$\psi_t = \eta (K \psi')', \tag{8}$$

where $\psi_t = (\partial / \partial t)\, \psi(V, t)$.

In sum, to lowest order the numerical code simulates the reduced plasma equations

$$\begin{aligned} \Delta \psi &= J(\psi, t), \\ \psi_t &= \eta (K \psi')', \\ \nabla \cdot \bar{U} &= 0, \end{aligned} \tag{9}$$

which is solved without reference to $\bar{U}$. The velocity has a higher order and could be found, if desired.

2. Alternating Dimension Numerical Formulation of the Model

The alternating dimension (AD or $1\frac{1}{2}$D) numerical formulation was used first in [5] and then in [6]. In the resistive case, considered here, the system of equations

(9) is separated into two equations which are solved alternately. The two equations are:

— the two-dimensional (2D) equilibrium equation,

$$\Delta\psi = J(\psi, t), \tag{10}$$

where $J(\psi, t)$ is assumed to be a known function of $\psi(x, y, t)$.

— the one-dimensional (1D) evolution equation for the poloidal magnetic flux, $\psi(V, t)$,

$$\psi_t(V, t) = \eta(K(V, t)\,\psi'(V, t))', \tag{11}$$

with $K(V, t)$ given. After (10) is solved, $K(V, t)$ is computed for the solution of (11) from the contours $\psi = \text{constant}$.

The 1D and 2D equations are coupled by

$$J(\psi, t) = \eta(K(V, t)\,\psi'(V, t))'. \tag{12}$$

The code alternates between an inner loop which solves (10) and uses the result in an outer loop which solves (11).

Inner loop (an iterative algorithm for solving (10)). Initially an approximation to the family of contours $\psi_0 = \text{constant}$ and a profile $J_0(\psi_0, t_0)$ as the right side in (10) are given. The algorithm solves a *nonlinear elliptic free boundary problem* with appropriate Dirichlet boundary conditions. Specifically, we are considering a plasma domain separated by a free boundary from a vacuum domain bounded externally by a perfect conductor. The condition on the free boundary is that the cross-sectional area of the plasma is constant, $V = V_{\text{plasma}}$.

Let $\psi_n(x, y)$ represent the flux function at the nth iteration and $V_n(x, y)$ the area inside the contour $\psi_n(x, y) = \text{constant}$. Without further comment, we treat V_n as a function of ψ_n (and t) or, inversely ψ_n as a function of V_n. We iterate as follows until a fixed point solution of (10) is found:

$$\begin{aligned} \psi_{n+1} = \Delta^{-1}(J_n(\psi_n, t_0)) \quad & \text{in } D \\ J_0(\psi_0, t_0) \quad & \text{given,} \\ \psi = \psi_B(t_0) \quad & \text{given on } \partial D, \\ V_{\text{plasma}} \quad & \text{fixed.} \end{aligned} \tag{13}$$

Here D is a rectangular domain inside which the vacuum field and plasma is contained. (See Fig. 1.)

The solution is characterized geometrically by tracing the contours $\psi = \text{constant}$ on a two-dimensional mesh and calculating $V(\psi)$ on a one-dimensional mesh.

The (2D) equilibrium solution gives the following physical quantities at time t_0: $\psi(x, y, t_0)$, $\psi(V, t_0)$, $J(x, y, t_0)$, $J(V, t_0)$, and $K(V, t_0)$. This completes the *inner loop*

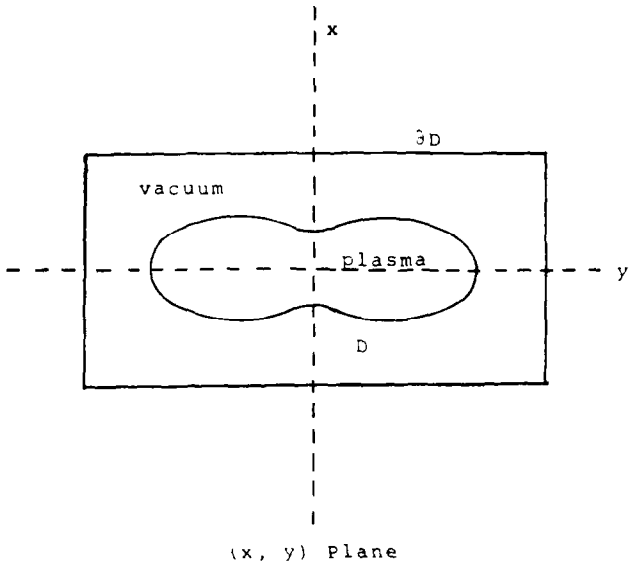

FIG 1 The domain, D, is the interior of the rectangle in the (x, y) plane The center of the rectangle is at the origin Dirichlet boundary conditions on ψ, ψ_B, are given on the perimeter of the rectangle, ∂D

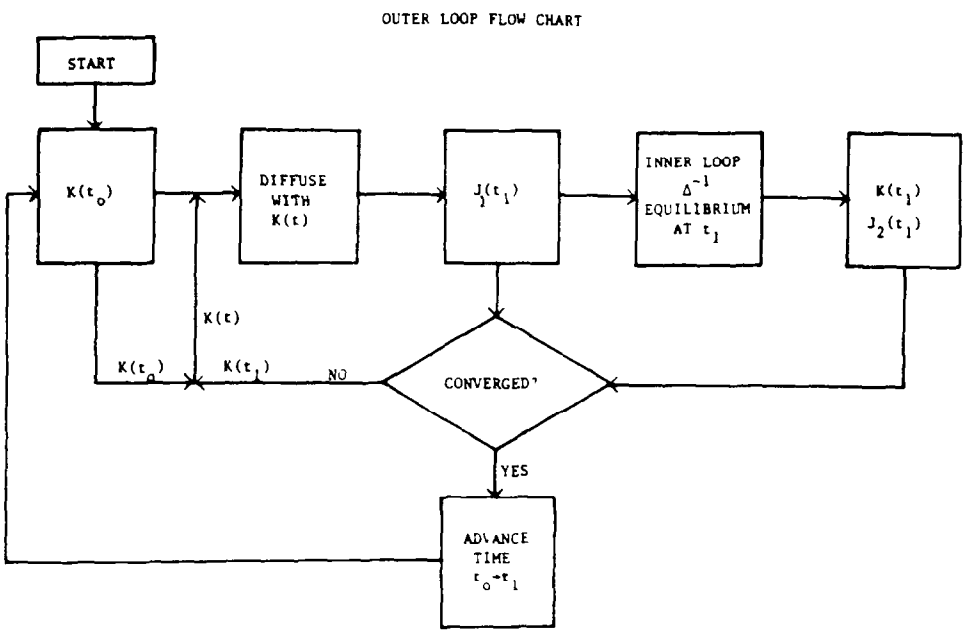

FIG 2 A flow chart of the *outer loop* iteration $K(t_0)$ and $K(t_1)$ are used to determine $K(t)$ numerically as piecewise linear function. Convergence is determined by comparing $J_1(t_1)$, the profiles resulting from the solution to the diffusion equation, to $J_2(t_1)$, the profiles resulting from the *inner loop* at time t_1

Outer loop. Using the equilibrium geometry at time t_0 obtained above, the 1D flux transport equation (Eq. (11)) is solved on a (one-dimensional) diffusion mesh for $t_0 \leqslant t \leqslant t_1$ with ψ given on the plasma edge.

The time dependence of $K(V, t)$ is numerically approximated by a piecewise linear function in t. This requires an outer loop Piecewise constant interpolation does not.

The outer loop (see Fig. 2), by alternating between the inner loop equilibrium solver and the 1D evolution solver, finds by iteration a self-consistent current profile $J(V, t_1)$, $t_1 > t_0$, which produces the quantities $\psi(V, t)$, $J(V, t)$, $K(V, t)$ for $t_0 \leqslant t \leqslant t_1$.

For surveys of AD ($1\frac{1}{2}$D) transport codes including the numerical formulation of many such codes and a detailed description of some problems which have been solved using them, see [11, 12].

3. The Doublet

A doublet is a special geometrical configuration of flux contours, $\psi = \text{constant}$, with a *separatrix*, a contour shaped like a figure eight, which bounds two *islands* of magnetic flux contours. (See Fig. 3.) It is used in [6] to study the evolution of the topology of contours. An interesting problem for numerical solution, which has important physical consequences, results from the formation and motion of the separatrix. This creates a very sharp peak in the current density profile. (See Fig. 4.)

The doublet is created by shaping coils, idealized in our case by specifying ψ on segments of the rectangular boundary (cf. (13)). We then simulated oscillating

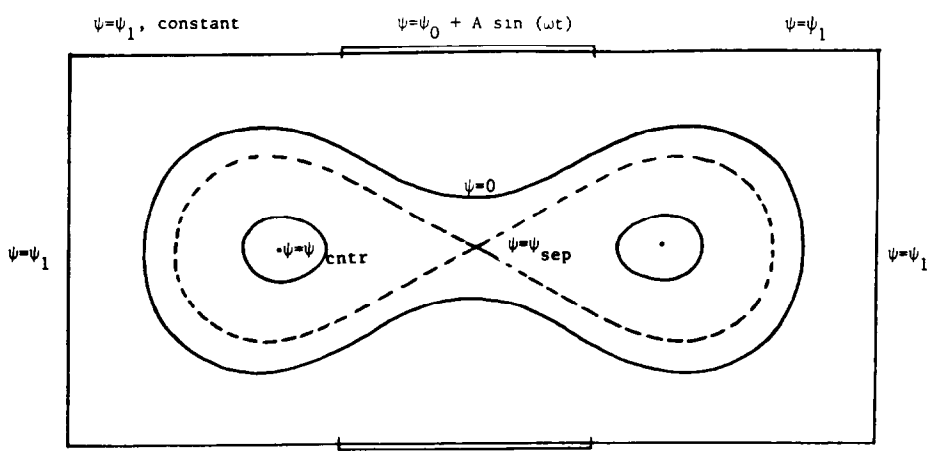

FIG 3 The doublet geometrical configuration A separatrix (dotted line) bounds two islands Lines represent $\psi = \text{constant}$ contours

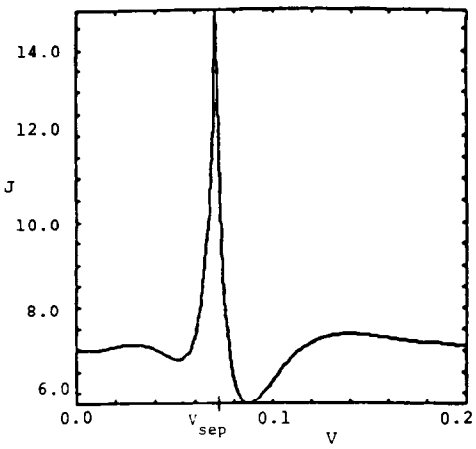

FIG 4 Current density profile A singularity in the form of a surface current on the separatrix arises in the doublet model of Grad *et al* [6] Our simulation indicates this by a sharp peak in the J-profile

currents in the shaping coils by superimposing a sinusoidally oscillating boundary condition on ψ in these segments and taking constant boundary values elsewhere on the boundary. This is a 2D analog (low β, high aspect ratio) of the Doublet III experiment at GA Technologies, Inc. [13] in the late '70s.

The doublet geometry is treated as a free boundary problem requiring the assignment of values of the volume,[1] V, and ψ to the separatrix. For this we proceed as follows: We introduce $V_{\text{sep}}(t)$, the volume of the separatrix, giving it a smooth piecewise cubic representation. Accordingly, the coordinates inside the separatrix are in terms of $V/V_{\text{sep}}(t)$ and outside $(V - V_{\text{sep}}(t))/(V_p - V_{\text{sep}}(t))$, where $V_{\text{sep}}(t)$ (the volume inside the separatrix) is piecewise cubic in t and V_p (the total plasma volume) is constant. The volume $V_{\text{sep}}(t)$ and its derivative $V'_{\text{sep}}(t)$ are approximated by continuous interpolation between t_0 and $t_0 + dt$. These interpolations are calculated by iteration in the outer loop. Because K is singular at the separatrix and J is highly peaked [6], normalization is particularly important since the singularity in normalized volume becomes stationary. The outer loop calculation is extremely sensitive to the position and velocity of the singularity. The iteration must therefore receive very special treatment; the details will be treated in Section 9.

Back-averaging will have three uses: to solve the nonlinear elliptic free boundary problem (inner loop), to determine the self-consistent J profiles (outer loop), and to calculate the velocity of the separatrix

[1] Area = volume/unit height

4 Back-Averaging and Fixed Point Iteration Schemes in General

Back-averaging is a multi-stage acceleration technique for calculating a fixed point $\bar{X}$ of a transformation $F: V \to V$ of a finite dimensional vector space V. Although the method applied to nonlinear problems, for the purpose of analysis and optimization of the back-average schemes, we later assume that F is approximately linear in the neighborhood of a fixed point.

We review the terminology and concepts for fixed point iteration schemes. A *fixed point* of F satisfies

$$F(\bar{X}) = \bar{X}. \tag{14}$$

The iteration scheme

$$\bar{X}_{N+1} = F(\bar{X}_N) \tag{15}$$

converges provided that the Lipschitz condition

$$\|F(P) - F(Q)\| \leqslant C\|P - Q\|, \qquad C < 1,$$

holds on a neighborhood of the fixed point which contains $\bar{X}_0$. If F is differentiable, instead of the Lipschitz condition, we shall use the bound on the derivative,

$$\|F'(\bar{X})\| < 1. \tag{16}$$

In general $\|\bar{X}\|$ denotes an appropriate norm of $\bar{X}$. Here we use the Euclidean norm but others are applicable and numerically useful.

Since F is differentiable we can approximate F by a linear function in the neighborhood of a fixed point $\bar{Y}$ by

$$F(\bar{X}) \approx L \cdot \bar{X} + \bar{B}, \tag{17}$$

where $L = F'(\bar{Y})$. In practice, we shall have only approximations to L and B from a priori calculations. We employ the iteration scheme for Eq. (17):

$$\begin{aligned} \bar{X}_{N+1} &= L \cdot \bar{X}_N + \bar{B}, \\ \bar{X}_0 &= \bar{B}. \end{aligned} \tag{18}$$

This yields $\bar{X}_N = (1 + L + L^2 + \cdots + L^N)\,\bar{X}_0$. The convergence of the iteration scheme in Eq. (18) is equivalent to that of the series:

$$(1 - L)^{-1} = 1 + L + L^2 + L^3 + \cdots. \tag{19}$$

For the sequence $\{\bar{X}_N\}$ defined by Eq. (18), we define

$$\Delta_N = \|\bar{X}_N - \bar{X}_{N-1}\|, \tag{20}$$

$$\lambda_N = \frac{\Delta_{N+1}}{\Delta_N}. \tag{21}$$

If $\{\bar{x}_N\}$ behaves like a geometric sequence, then λ_N has a limit λ. This λ will be the absolute value of the largest eigenvalue of L. Convergence is guaranteed if $|\lambda| < 1$ but our problem does not always satisfy that condition. In fact, $|\lambda|$ less than but close to 1 is not useful. (Even in our "easy" case $|\lambda| = 0.9979$.) We need a method of obtaining convergence when it does not exist or accelerating it when it does. That is the topic of the next section.

5. Back-Averaging

For the purpose of generating or accelerating convergence, Stevens *et al.*, used the method of simple back-averaging (SBA) [5], also called two-point back-averaging. Here, we furnish an analysis and optimization of SBA and extend the method to a more powerful technique, three-point back-averaging. Simple back-averaging is analogous to the method of successive overrelaxation (SOR) [14].

Consider a transformation from the sequence $\{X_N\}$ of (18) to a new sequence $\{Y_N\}$ defined as

$$\begin{aligned} &\{\bar{X}_N\} \to \{\bar{Y}_N\}, \\ &\quad \bar{Y}_0 = \bar{X}_0, \\ &\bar{Y}_{N+1} = (1-\alpha)\, F(\bar{Y}_N) + \alpha \bar{Y}_N, \qquad \alpha \in \mathbf{R}. \end{aligned} \tag{22}$$

If $\{X_N\}$ is governed by the affine transformation $T: \bar{X} \to L\bar{X} + \bar{B}$, Eq. (18), then $\{Y_N\}$ will be defined by

$$\begin{aligned} Y_{N+1} &= \tilde{T} Y_N, \\ \tilde{T} &= (1-\alpha) T + \alpha \end{aligned} \tag{23}$$

If $\{\bar{X}_N\}$ becomes geometric, i.e., $\bar{X}_{N+1} = \lambda \bar{X}_N$, λ complex and constant for large N, then

$$\begin{aligned} \bar{Y}_{N+1} &= \tilde{\lambda} \bar{Y}_N, \\ \tilde{\lambda} &= (1-\alpha)\lambda + \alpha. \end{aligned} \tag{24}$$

Two-point back-averaging is so called because two successive stages $\bar{Y}_N$ and $F(\bar{Y}_N)$ are used at each iteration step. The concept of SBA is simple: the nth iterate Y_{N+1} is not $F(Y_N)$ alone, but a linear combination of Y_N and $F(Y_N)$. In the simplest form Y_{N+1} is a weighted average of Y_N and $F(Y_N)$ with weights between zero and one. More generally, the weights are less restricted; we require only that they are real. This increase in generality is crucial.

Geometrically, we interpret the connection between $\tilde{\lambda}$ and λ for a given α as follows. Consider $\lambda = \xi + i\eta$ in the complex plane. Draw a line (L_1) passing through λ and (1, 0). Depending on the choice of α, $\tilde{\lambda}$ can lie anywhere on this line. If $\alpha = 0$ then $\tilde{\lambda} = \lambda$, and the iteration does not involve back-averaging. If $\alpha = 1$ then $\tilde{\lambda} = 1$ and the scheme is marginally convergent. Notice that α can be negative or > 1.

Next, consider two successive applications of SBA using a different back-average parameter at each application or stage. Using the linearized form of F, T, let

$$\begin{aligned} \tilde{T}_1 \circ \bar{X} &= (\alpha_1 + (1-\alpha_1)T) \circ \bar{X}, \\ \tilde{T}_2 \circ \bar{X} &= (\alpha_2 + (1-\alpha_2)T) \circ \bar{X}. \end{aligned} \tag{25}$$

Combining $\tilde{T}_1$ and $\tilde{T}_2$ gives the three-point back-average (3PBA) formula,

$$\tilde{T}_{\alpha_2\alpha_1} \circ \bar{X} = \tilde{T}_2 \circ \tilde{T}_1 \circ \bar{X} = (\alpha_2 + (1-\alpha_2)T)(\alpha_1 + (1-\alpha_1)T) \circ \bar{X}. \tag{26}$$

Using the definition of $\{\bar{X}_n\}$, Eq. (18), the 3PBA transformation is

$$\begin{aligned} \{\bar{X}_n\} &\to \{\bar{Y}_n\}, \\ \bar{Y}_1 &= \bar{X}_1, \\ \bar{Y}_{n+1} &= \alpha_1\alpha_2 \bar{Y}_n + [(1-\alpha_1)\alpha_2 + (1-\alpha_2)\alpha_1] T \circ \bar{Y}_n + (1-\alpha_1)(1-\alpha_2) T^2 \circ \bar{Y}_n. \end{aligned} \tag{27}$$

A three-point back-average iteration, $\tilde{T}_{\alpha_2\alpha_1}$, can be written as a 3-stage iteration scheme:

$$\begin{aligned} \bar{Y}_n^{(0)} &= \bar{Y}_n \\ \bar{Y}_n^{(1)} &= \alpha_1 \bar{Y}_n^{(0)} + (1-\alpha_1) T \circ \bar{Y}_n^{(0)} \\ \bar{Y}_n^{(2)} &= \alpha_2 \bar{Y}_n^{(1)} + (1-\alpha_2) T \circ \bar{Y}_n^{(1)} \\ \bar{Y}_{n+1} &= \bar{Y}_n^{(2)} \end{aligned} \tag{28}$$

Thus 3PBA is a multi-stage iteration scheme [15].

We shall describe an optimal choice of α_1 and α_2. Earlier in the literature of plasma equilibria, Marder and Weitzner [16] gave a special 3-stage iterative method which can be written as a 3PBA scheme but is not optimal. This is given by

$$\bar{Y}_{n+1} = (1-\beta^2)\, \bar{Y}_n + 2\beta^2 F(\bar{Y}_n) - \beta^2 F(F(\bar{Y}_n)), \tag{29}$$

In the linear case, this scheme can be factored as

$$\bar{Y}_{n+1} = [(1-\beta) + \beta T)][(1+\beta) - \beta T] \circ \bar{Y}_n, \tag{30}$$

which amounts to (27) with

$$\begin{aligned} \alpha_1 &= (1-\beta), \\ \alpha_2 &= (1+\beta). \end{aligned} \tag{31}$$

Unfortunately, in our application MWS tends to converge slowly. To overcome this slow rate of convergence, at the expense (or benefit) of changing the region of convergence and perhaps shrinking it, we apply 3PBA by treating α_1 and α_2 as independent parameters.

6. Region of Convergence

Recall that in two-point back-averaging λ is the eigenvalue of T and $\tilde{\lambda}$ the corresponding eigenvalue of $\tilde{T}$, see (24).

The image of $|\tilde{\lambda}| < 1$ in the complex λ-plane, is the interior of the circle of marginal convergence CM_α, defined by

$$\left|\lambda - \frac{\alpha}{\alpha - 1}\right| = \frac{1}{|\alpha - 1|}, \tag{32}$$

with center at $\alpha/(\alpha - 1)$ on the real axis and passing through the point $\lambda = 1$. If $\tilde{T}$ is iterated with this value of α, all eigenvalues λ inside CM_α yield convergence.

If $\mathrm{Re}\{\lambda\} > 1$, $\{\bar{Y}_N\}_\alpha$, the sequence obtained by back-averaging with α, can be made to converge by taking $\alpha = 1 + \varepsilon$ with sufficiently small positive ε. Given $\mathrm{Re}\{\lambda\} < 1$, $\{\bar{Y}_N\}_\alpha$ converges for $\alpha = 1 - \varepsilon$ and sufficiently small positive ε (see Fig. 5).

If the eigenvalues can be enclosed in a circle CM_α, $\{\bar{Y}_N\}_\alpha$ converges. Therefore, when all eigenvalues have $\mathrm{Re}\{\lambda\} < 1$, an α can always be found which yields con-

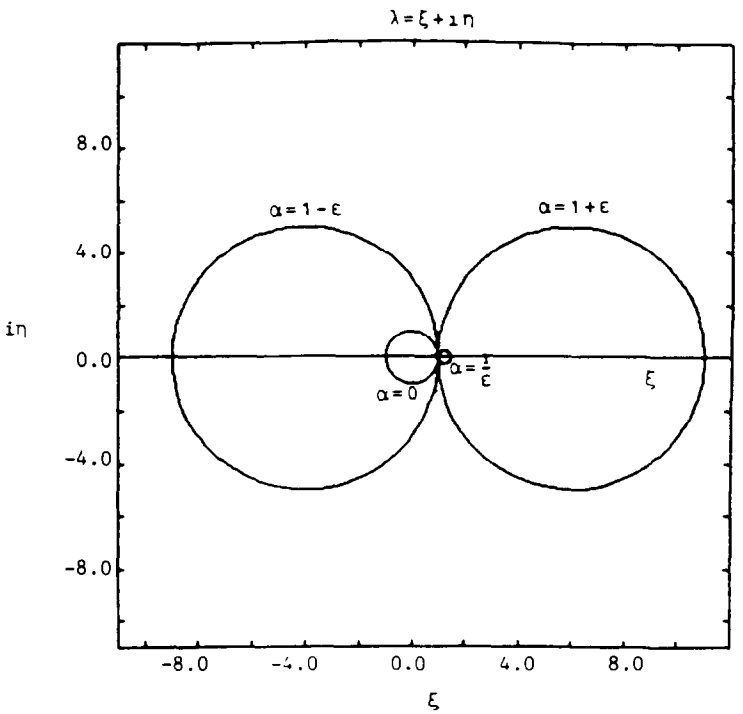

FIG 5. Circles of marginal convergences CM_α If λ is an eigenvalue of T and lies inside CM_α then the calculation of the eigenfunction corresponding to the eigenvalue λ converges Five curves CM_α are shown The largest two are the cases $\alpha = 1 - \varepsilon$ and $\alpha = 1 + \varepsilon$ Next in size is a $\alpha = 0$ (no back averaging), next is $\alpha = 1/\varepsilon$ In all cases, $\varepsilon = 0.2$ The dotted line is $\xi = 1$ which is tangent to CM_α for all α at $\eta = 0$

vergence. Likewise, when all eigenvalues have $\mathrm{Re}\{\lambda\} > 1$, an α can be found which yields convergence. However, no simple back-average scheme converges if there are two eigenvalues, λ_1 and λ_2, satisfying $\mathrm{Re}\,\lambda_1 > 1$ and $\mathrm{Re}\,\lambda_2 < 1$. In addition, if $\lambda = 1 + i\eta$ for any η, no simple back-average scheme converges.

For a three-point back-averaging scheme, the region of convergence $|\tilde{\lambda}| = 1$ of $\tilde{T}_{\alpha_1\alpha_2}$ is calculated as

$$|\tilde{\lambda}| = |\alpha_2 + (1-\alpha_2)\lambda|\,|\alpha_1 + (1-\alpha_1)\lambda|. \tag{33}$$

Equivalently for $\rho = |\tilde{\lambda}|$,

$$\begin{aligned}\rho^2 = \{&\alpha_1\alpha_2 + [(\alpha_1+\alpha_2) - 2\alpha_1\alpha_2]\xi + [1 + \alpha_1\alpha_2 - (\alpha_1+\alpha_2)](\xi^2 - \eta^2)\}^2 \\ &+ \eta^2\{(\alpha_1+\alpha_2) - 2\alpha_1\alpha_2 + 2[1 + \alpha_1\alpha_2 - (\alpha_1+\alpha_2)]\xi\}^2.\end{aligned} \tag{34}$$

In the above, substitute

$$\begin{aligned}\alpha_1 + \alpha_2 &= 2 - \delta, \\ \alpha_1\alpha_2 &= 1 - \varepsilon,\end{aligned} \tag{35}$$

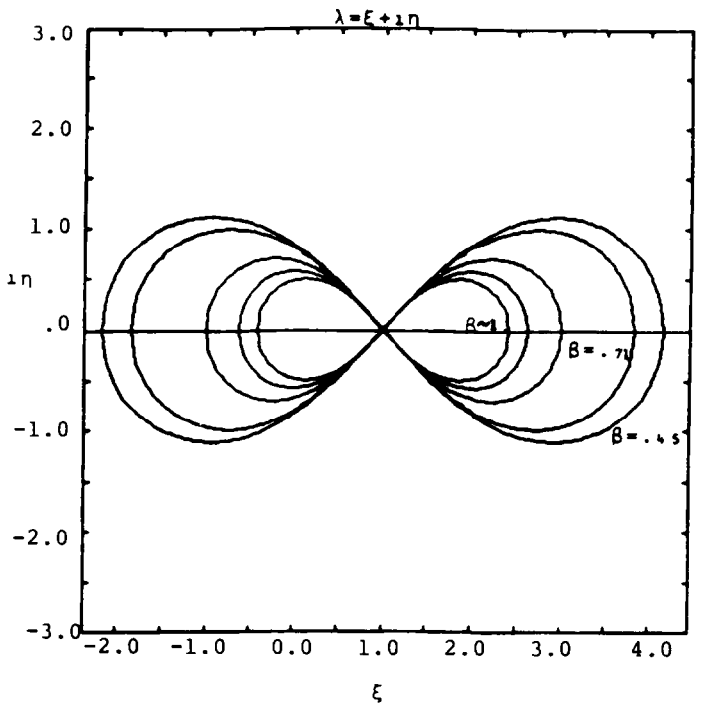

FIG 6 Curves of marginal convergence for the Marder–Weitzner iteration scheme are lemniscates L_β The crossing point is $\xi = 1$ and the petals are asymptotic to $\xi = \eta$ and $\xi = -\eta$ As β is decreased L_β becomes larger In the above the largest curve corresponds to $\beta = 0\,45$, the next smaller to $\beta = 0.5$, then $\beta = 0\,71$, $\beta = 0\,87$ and the smallest, $\beta = 1.00$

which gives

$$\rho^2 = [(1-\varepsilon) + (2\varepsilon - \delta)\xi + (\delta - \varepsilon)(\xi^2 - \eta^2)]^2 + \eta^2[2\varepsilon - \delta + 2(\delta - \varepsilon)\xi]^2. \quad (36)$$

In (36), we set

$$\tilde{\xi} = \frac{1}{\sigma}(\delta - \varepsilon)^2 \left[\xi + \frac{1}{2}\left(\frac{2\varepsilon - \delta}{\delta - \varepsilon}\right)\right]^2,$$

$$\tilde{\eta} = \frac{1}{\sigma}(\delta - \varepsilon)^2 \eta^2, \quad (37)$$

$$\tilde{\rho} = \frac{1}{\sigma^2}(\delta - \varepsilon)^2 \rho^2,$$

$$\sigma = \frac{\delta^2}{4} - (\delta - \varepsilon) = \frac{1}{4}(\alpha_1 - \alpha_2)^2 > 0. \quad (38)$$

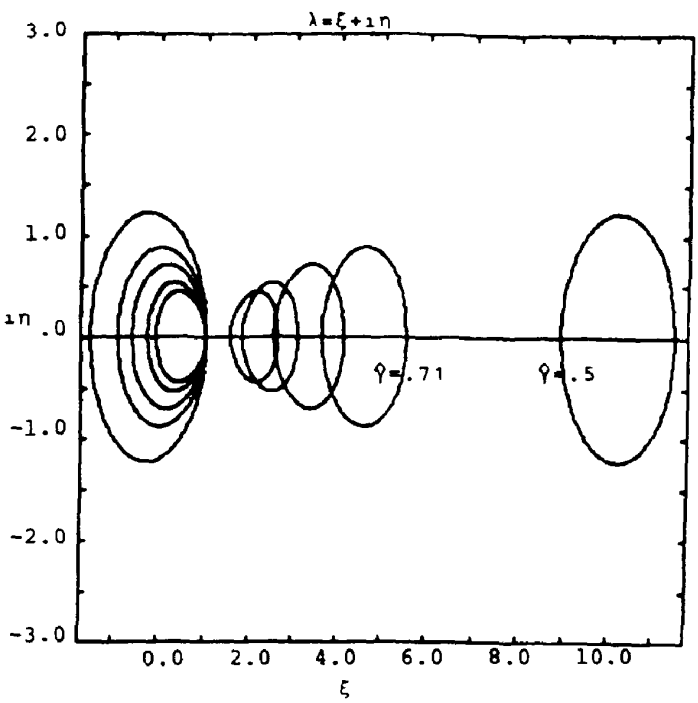

FIG. 7. Curves of marginal convergence, $\Gamma_{\hat{\gamma}\beta}$ These are disjoint ovals The center of symmetry ξ_0 shifts to the right of $\xi = 1$ and the size of the ovals increases as $\hat{\gamma}$ is decreased. Notice that all the ovals on the left side of $\xi = 1$ pass through (1, 0) In this figure, the curves $\Gamma_{\hat{\gamma}\beta}$ are shown for the following $(\hat{\gamma}, \beta)$ pairs. (0 5, 1.8), (0.71, 1 8), (0 87, 1 8), (1 , 1 8), and (1 22, 1.8)

We get

$$\tilde{\xi}^2 + \tilde{\eta}^2 + 2\tilde{\xi}\tilde{\eta} - 2(\tilde{\xi} - \tilde{\eta}) + (1 - \tilde{\rho}) = 0 \tag{39}$$

and hence

$$\tilde{\eta} = -(\tilde{\xi} + 1) \pm [(\tilde{\xi} + 1)^2 - (\tilde{\xi}^2 - 2\tilde{\xi} - 1 - \tilde{\rho}]^{1/2}. \tag{40}$$

For any α_1, α_2 the contours $\rho = \text{constant}$ belong to the family of curves in the ξ, η-plane consisting of ovals of Cassini, lemniscates, and "peanuts." In particular, the *curve of marginal convergence*, $\rho = 1$, which bounds the set of eigenvalues for which 3PBA converges, is a member of this family.

Let the curve of marginal convergence be denoted by $\Gamma_{\alpha_1\alpha_2}$ or, equivalently, in terms of ε and δ, by $\Gamma_{\varepsilon\delta}$. Recall from Section 5 that MWS is 3PBA with $\alpha_1 + \alpha_2 = 2$, $\alpha_1\alpha_2 = 1 - \beta^2$. In terms of ε and δ, cf. (31) and (35), MWS is equivalent to 3PBA with $\delta = 0$ and $\varepsilon = \beta^2$. In this case, the curve of marginal convergence $\Gamma_{\varepsilon 0}$ is a lemniscate L_β. (See Fig. 6.)

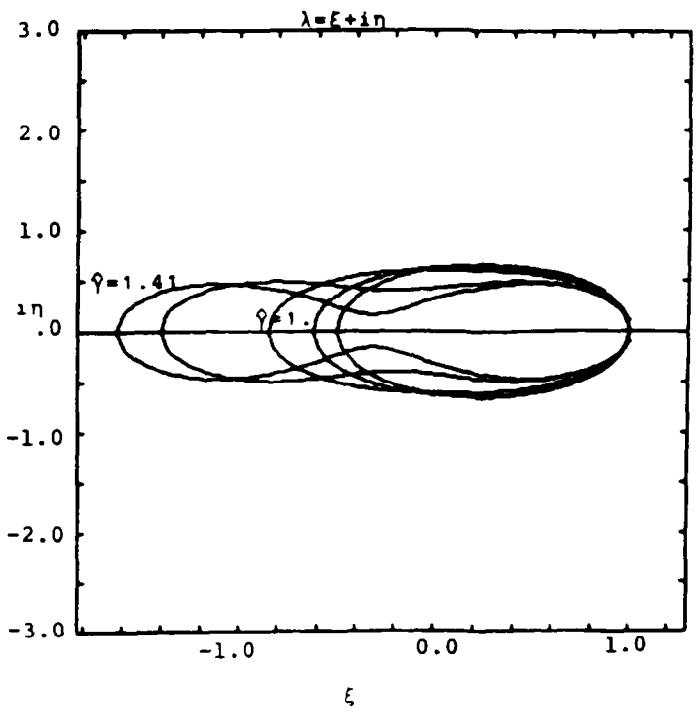

FIG 8 The degenerate case in which there are only two $\eta = 0$ intercepts. $\Gamma_{\beta\beta}$ are single ovals or peanuts. $\beta = -1$ in all the above, $\hat{\gamma} = 1.41$ for the largest peanut, $\hat{\gamma} = 0.71$ for the smallest oval. The intermediate curves are $\hat{\gamma} = 1.22$, $\hat{\gamma} = 1.00$, and $\hat{\gamma} = 0.87$.

We replace α_1, α_2 by the parameters

$$\begin{aligned} \hat{\beta} &= \alpha_1 + \alpha_2, \\ \hat{\gamma} &= \alpha_1 \alpha_2, \end{aligned} \tag{41}$$

to depict the family in the different parts of the parameter range (see Figs. 7 and 8.)

7. Choosing an Optimal Simple Back-Average Parameter

We seek an α for which the iteration of $\tilde{T}$ converges quickly. Let e be the eigenfunction of T with eigenvalue λ ($\mathrm{Re}(\lambda) \neq 1$). Set $\rho = |\tilde{\lambda}|$ for $\tilde{\lambda}$ corresponding to the specific eigenvalue, $\lambda = \xi + i\eta$ of T, and back-average parameter α. We have

$$\rho^2 = |\tilde{\lambda}|^2 = [(1-\alpha)^2(\xi^2 + \eta^2) + 2\alpha(1-\alpha)\xi + \alpha^2]. \tag{42}$$

The convergence to the eigenfunction will be optimal for the value α_* of α which yields the minimum value ρ_*^2 of ρ^2.

$$\alpha_* = \frac{\eta^2 + \xi(\xi - 1)}{\eta^2 + (\xi - 1)^2}. \tag{43}$$

With this choice of α_*, for $\xi \neq 1$, convergence is assured:

$$\rho_*^2 = \min |\tilde{\lambda}|^2 = \frac{\eta^2}{\eta^2 + (\xi - 1)^2} < 1. \tag{44}$$

If λ is real, then there exists an α_* such that $\rho_* = 0$. This will sum the series in one iteration and eliminate the eigenfunction belonging to λ. In this case,

$$\alpha_* = \frac{\lambda}{\lambda - 1} \tag{45}$$

Note that the optimum back-average parameter α_* depends on the value of the eigenvalue λ. If T has a dominant eigenvalue, then the convergence of the iteration scheme may be satisfactory if the back-average α_* is calculated using the dominant eigenvalue. However, if the spectrum of T is broad, convergence can be improved by taking an optimum α_{opt} of α relative to a set, S, of N eigenvalues, $S = \{\lambda_j = \xi_j + i\eta_j, j = 1, ..., N\}$.[2]

Finding α_{opt} is a nonlinear minimax optimization problem. Algorithms for solving the general nonlinear minimax optimization problem are given by Overton [17]. Here, we solve the optimization problem when S is a set of only *three* eigenvalues.

[2] The difficulty is that a lesser eigenvalue may emerge, after simple back-averaging, with an absolute value even greater than that of the originally largest eigenvalue

Consider the following minimax optimization problem for defining α_{opt},

$$\min_{\alpha} \{F_M(\alpha) \mid \alpha \in R\}, \tag{46}$$

where

$$F_M(\alpha) = \max\{\rho_j^2(\alpha),\ j = 1, 2, ..., N\}, \tag{47}$$

and

$$\rho_j^2(\alpha) = |\tilde{\lambda}_j(\alpha)|^2; \tag{48}$$

here $\tilde{\lambda}_j(\alpha)$ is the eigenvalue of $\tilde{T}$ from back-averaging with α which corresponds to the eigenvalue λ of T.

Notice that the ρ_j are convex functions of α. F_M is also convex. Since F_M is

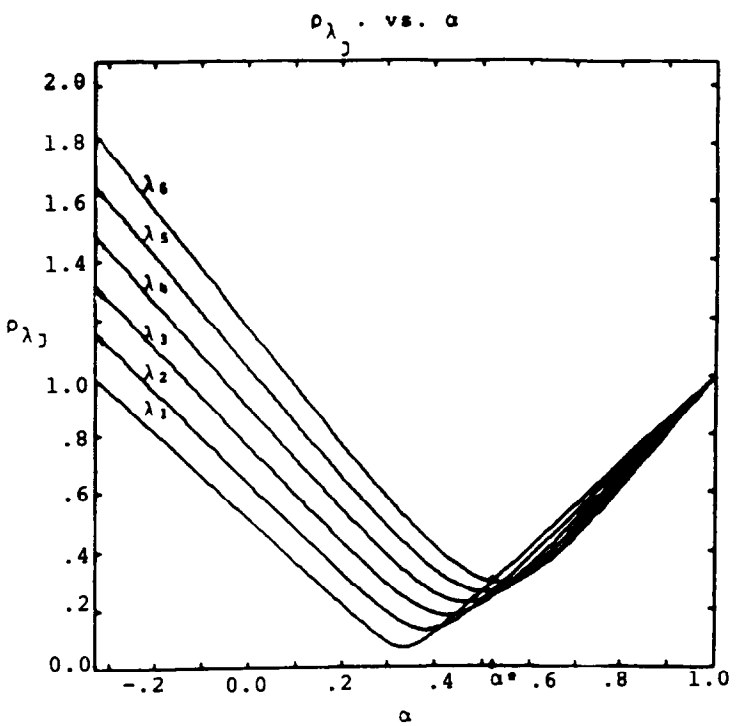

FIG 9. Curves $\rho_{\lambda_j}(\alpha)$ for $\lambda_j = \xi_j + i\eta_j$ The curves are convex with minimum at $\alpha^*(\lambda_j)$ given by Eq (44) $\rho_{\lambda_j}(\alpha)$ is never zero for $\text{Im}(\lambda_j) \neq 0$, however, $\rho_{\lambda_j}(\alpha^*(\lambda_j)) < 1$ for all λ_j. If only one λ_j is considered in determing the optimum then $\alpha^*(\lambda_j)$ is the result If more than one λ_j is considered, the minimax optimization algorithm outlined in Section 7 determines α^* In this case the intersection of ρ_{λ_1} and ρ_{λ_6} gives α^* In the above $\lambda_1 = 0\,5 + 0\,1i$, $\lambda_2 = -0\,6 + 0.2i$, $\lambda_3 = -0.7 + 0\,3i$, $\lambda_4 = -0\,8 + 0.4i$, $\lambda_5 = -0\,9 + 0\,5i$, $\lambda_6 = 1 + 0\,6i$

continuous and piecewise differentiable, each piece being a segment of some ρ_j, the global minimum of F_M will occur either at a minimum of $\rho_j(\alpha)$ for some j or at an intersection of $\rho_j(\alpha)$ with $\rho_k(\alpha)$ for some (j, k). (See Fig. 9.)

With $N=3$, an optimization algorithm need not be sophisticated. We shall use the following optimization procedure:

1. We find the intersections, $\alpha_{\text{int}}(\lambda_j, \lambda_k)$ of $\rho_j(\alpha)$ and $\rho_k(\alpha)$, for all (λ_j, λ_k) pairs $j, k = 1, 2, 3$ using Eq. (49), namely,

$$\alpha_{\text{int}}(\lambda_j, \lambda_k) = 1 \Big/ \left\{ 1 - \left[\frac{2(\text{Re}\,\lambda_j - \text{Re}\,\lambda_k)}{|\lambda_j|^2 - |\lambda_k|^2} \right] \right\}. \tag{49}$$

and form

$$\text{INT} = \{\alpha_{\text{int}}(\lambda_j, \lambda_k) < 1,\ \text{pairs}(\lambda_j, \lambda_k),\ j \neq k,\ j, k = 1, 2, 3\}. \tag{50}$$

(Besides the intersection given by (49), there is a second intersection at $\alpha = 1$, where $\rho(1) = 1$ for all eigenvalues.)

2. Let λ_{mx} yield the maximum of the minima of the several $\rho_j(\alpha)$; λ_{mx} satisfies

$$|\lambda_{\text{mx}} - 1| = \max_{\lambda_j \in S} \left(\frac{\eta_j^2}{|\lambda_j - 1|^2} \right). \tag{51}$$

3. If the set INT of intersections is empty, then $\alpha_{\text{opt}} = \alpha_*(\lambda_{\text{mx}})$ as defined by Eq. (43).

4. If $\text{INT} \neq \{\phi\}$ then set $\text{TEST} = \text{INT} \cup \{\alpha_*(\lambda_j), j = 1, 2, 3\}$.

5. From TEST form the set TEST′ by deleting any elements which yield points below the graph of F_M; that is, delete all α_j^* for which $\rho_j(\alpha_j^*) < F_M(\alpha_j^*)$ and all α_{ij} for which $\rho_i(\alpha_{ij}) < F_M(\alpha_{ij})$, where $\alpha_{ij} = \alpha_{\text{int}}(\lambda_i, \lambda_j)$.

6. In conclusion, α_{OPT} is given by

$$\rho(\alpha_{\text{OPT}}) = \min_{\alpha \in \text{TEST}'} \{F(\alpha)\}. \tag{52}$$

The cost of choosing α_{opt} based on three eigenvalues is negligible. However, if a significant number of eigenvalues are used, the cost would become prohibitive. The cost of calculating N eigenvalues involves solving an $N \times N$ inverse problem and $N+1$ iterations of F to find T after which the eigenvalue problem must be solved. The additional information gained by using a large N is not cost effective because the original transformation F is nonlinear and T is only an approximation, which, at times is quite noisy. Therefore, the benefit of an extremely precise determination of α will be counteracted by the noise in the system. In practice either three or four eigenvalues are calculated. However, only the two most dominant eigenvalues are used. We assume the third eigenvalue is zero to accommodate the many eigenvalues near zero.

8. Optimization of Three-Point Back-Averaging

For 3PBA, there are two back-average parameters, α_1 and α_2, available for optimization. Denote the optimal pair of back-average parameters by (α_1^*, α_2^*).

In Section 7 for SBA, using one eigenvalue λ to determine the optimum, we obtain the optimal α_* from Eq. (43). For 3PBA with only one eigenvalue, λ, the optimal pair becomes (α_*, α_*), where α_* satisfies the same equation. This follows immediately from the following: From Eq. (26),

$$\|\tilde{T}_{\alpha_1,\alpha_2}\| = \|\alpha_1 + (1-\alpha_1)T\| \, \|\alpha_2 + (1-\alpha_2)T\|. \tag{53}$$

Consequently, when we use only one eigenvalue λ,

$$\min_{\alpha_1\alpha_2} \|\tilde{\lambda}_{\alpha_1\alpha_2}\| = \min_{\alpha_1} \|\alpha_1 + (1-\alpha_1)\lambda\| \min_{\alpha_2} \|\alpha_2 + (1-\alpha_2)\lambda\|. \tag{54}$$

α_* satisfying (43) minimizes each factor in (54); hence, the optimal pair is (α_*, α_*).

If two eigenvalues, λ_1 and λ_2, are used to determine the optimal pair (α_1, α_2) and each α is real, then the optimal pair is $(\alpha_*(\lambda_1), \alpha_*(\lambda_2))$, where

$$\alpha_*(\lambda_j) = \frac{\lambda_j}{\lambda_j - 1}. \tag{56}$$

This is a consequence of the following proposition:

> Given a linear transformation T of a linear vector space V of dimension N. Assume that T has exactly M distinct eigenvectors, $\bar{e}_j$, whose eigenvalues are λ_j, $j = 1, 2, 3, ..., M$. If λ_j are real for all j, then there exist α_j^*, $j = 1, 2, ..., M$ such that the Euclidean norm, $\|\tilde{T}_{\alpha_1^*\alpha_2^*\ \ \alpha_M^*}\|$, of the $M+1$

We omit the proof which is straightforward.

For real λ_1 and λ_2, α_1^* and α_2^* are readily determined. For complex eigenvalues $\lambda_j = \xi_j + i\eta_j$, $j = 1, 2$, if $\rho_{\lambda_1}(\alpha_1, \alpha_2) \neq \rho_{\lambda_2}(\alpha_1, \alpha_2)$ for any (α_1, α_2) other than $(1, 1)$, the optimal (α_1^*, α_2^*) will be the pair which minimizes the maximum of $\rho_{\lambda_1}(\alpha_1, \alpha_2)$ and $\rho_{\lambda_2}(\alpha_1, \alpha_2)$. If $\rho_{\lambda_1}(\alpha_1, \alpha_2) = \rho_{\lambda_2}(\alpha_1, \alpha_2)$ along a curve, Ω, in the (α_1, α_2) plane the optimum occurs at a point on Ω for which $\rho_{\lambda_1}(\alpha_1, \alpha_2)$ is minimum or elsewhere at a minimum of $\rho_{\lambda_1}(\alpha_1, \alpha_2)$ or $\rho_{\lambda_2}(\alpha_1, \alpha_2)$. The extremum on Ω can be calculated by minimizing the Lagrange multiplier function, $\mathscr{L}$:

$$\begin{aligned}\mathscr{L} = {} & |\tilde{\alpha}\lambda_1 + (1-\tilde{\alpha})| \, |\tilde{\beta}\lambda_1 + (1-\tilde{\beta})| + \mu[|\lambda_2\tilde{\alpha} + (1-\tilde{\alpha})| \, |\lambda_2\tilde{\beta} + (1-\tilde{\beta})| \\ & - |\lambda_1\tilde{\alpha} + (1-\tilde{\alpha})| \, |\lambda_1\tilde{\beta} + (1-\tilde{\beta})|],\end{aligned} \tag{57}$$

where $\tilde{\alpha} = 1 - \alpha_1$ and $\tilde{\beta} = 1 - \alpha_2$. In terms of $\mathscr{L}$ the extremum problem becomes

$$\min_{\tilde{\alpha}, \tilde{\beta}, \mu} [\mathscr{L}], \tag{58}$$

subject to

$$|\lambda_2\tilde{\alpha}+(1-\tilde{\alpha})|\,|\lambda_2\tilde{\beta}+(1-\tilde{\beta})|-|\lambda_1\tilde{\alpha}+(1-\tilde{\alpha})|\,|\lambda_1\tilde{\beta}+(1-\tilde{\beta})|=0.$$

Since Ω is convex, we determine the minimum on Ω as the solution of

$$\frac{\partial\mathscr{L}}{\partial\tilde{\alpha}}=\frac{\partial\mathscr{L}}{\partial\tilde{\beta}}=\frac{\partial\mathscr{L}}{\partial\mu}=0. \tag{59}$$

Using $\hat{\lambda}_1=1-\lambda_1$ and $\hat{\lambda}_2=1-\lambda_2$, yields the system

$$\frac{(|\hat{\lambda}_2|^2\,\tilde{\alpha}-\operatorname{Re}\hat{\lambda}_2)(|\hat{\lambda}_1|^2\,\tilde{\beta}-\operatorname{Re}\hat{\lambda}_1)}{(|\hat{\lambda}_1|\,\tilde{\alpha}-\operatorname{Re}\hat{\lambda}_1)(|\hat{\lambda}_2|\,\tilde{\beta}-\operatorname{Re}\hat{\lambda}_2)}=\frac{|\hat{\lambda}\tilde{\beta}-1|^4}{|\hat{\lambda}_2\tilde{\beta}-1|^4}, \tag{60}$$

$$\tilde{\gamma}^2-2\tilde{\gamma}\left[\frac{|\hat{\lambda}_1|^2-|\hat{\lambda}_2|^2+\tilde{\delta}(|\hat{\lambda}_1|^2\operatorname{Re}\hat{\lambda}_1-|\hat{\lambda}_2|^2\operatorname{Re}\hat{\lambda}_2)}{|\hat{\lambda}_1|^4-|\hat{\lambda}_2|^4}\right]$$
$$+\left[\frac{\tilde{\delta}^2(|\hat{\lambda}_1|^2-|\hat{\lambda}_2|^2-2\tilde{\delta}(\operatorname{Re}\hat{\lambda}_1-\operatorname{Re}\hat{\lambda}_2)}{|\hat{\lambda}_1|^4-|\hat{\lambda}_2|^4}\right]=0, \tag{61}$$

where

$$\tilde{\gamma}=\tilde{\alpha}\tilde{\beta}\qquad\text{and}\qquad\tilde{\delta}=\tilde{\alpha}+\tilde{\beta}.$$

The minimum problem is reduced to the solution of the 4th-order system (60) and (61) for $\tilde{\gamma}$ and $\tilde{\delta}$ which is accomplished numerically. Now we wish to solve the minimax problem corresponding to the one of Section 7. For this general optimization problem, a set, S, of N eigenvalues, λ_j, determines the optimum pair, $(\alpha_1^{**}, \alpha_2^{**})$. Hence let $(\alpha_1^{**}, \alpha_2^{**})$ be the pair which solves the following problem:

$$\min_{(\alpha_1\ \alpha_2)}\{F_M(\alpha_1,\alpha_2)\,|\,(\alpha_1,\alpha_2)\in\mathbf{R}^2\}, \tag{62}$$

where

$$F_M(\alpha_1,\alpha_2)=\max_{\lambda_j\in S}\{\rho_{\lambda_j}(\alpha_1,\alpha_2)\} \tag{63}$$

and

$$\rho_{\lambda_j}(\alpha_1,\alpha_2)=\|\tilde{\lambda}_j\|. \tag{64}$$

Here $\tilde{\lambda}_j$ is the eigenvalue of $\tilde{T}_{\alpha_1\alpha_2}$ corresponding to the eigenvalue λ_j of T.

Notice that $\rho=\rho_{\lambda_j}(\alpha_1,\alpha_2)$, $j=1,2,\ldots$ represent convex surfaces in three-space (ρ,α_1,α_2). Moreover, F_M is composed of sections of $\rho_{\lambda_j}(\alpha_1,\alpha_2)$ surfaces for some $\lambda_j\in S$ and is convex. The global solution to the minimax problem lies at a minimum of $\rho_{\lambda_j}(\alpha_1,\alpha_2)$ or at an intersection of $\rho_{\lambda_j}(\alpha_1,\alpha_2)$ with $\rho_{\lambda_k}(\alpha_1,\alpha_2)$ for some λ_j, λ_k. For simple back-averaging, α represented the only degree of freedom and the intersec-

tions are points. Now there are two degrees of freedom α_1, α_2 and the intersections are curves

The optimization concepts can be generalized to any finite number, K, of degrees of freedom, $\alpha_1, \alpha_2, ..., \alpha_K$.

For the current application we use three eigenvalues to determine the optimum. The explicit procedure for determining the optimum pair $(\alpha_1^{**}, \alpha_2^{**})$ follows:

1. Calculate the optimum pair, $(\alpha_1^*(\lambda_j, \lambda_k), \alpha_2^*(\lambda_j, \lambda_k))$, for each pair of eigenvalues (λ_j, λ_k), $j \neq k = 1, 2, 3$, using the Lagrangian multiplier (58).

2. Form the set OP of optimum pairs,

$$\mathrm{OP} = \{(\alpha_1^*(\lambda_j, \lambda_k), \alpha_2^*(\lambda_j, \lambda_k)), j \neq k = 1, 2, 3\} \tag{65}$$

considering only two eigenvalues λ_j and λ_k. (Note that this could be a minimum of ρ_{λ_j} or ρ_{λ_k}, or a minimum within the intersection, Ω.)

3. Form the set OQ where

$$\mathrm{OQ} = \{\rho_l^*(\lambda_j, \lambda_k), l = 1, 2, 3\}, \tag{66}$$

where

$$\rho_l^*(\lambda_j, \lambda_k) = \max_m \{\rho_{\lambda_m}(\alpha_1^*(\lambda_j, \lambda_k), \alpha_2^*(\lambda_j, \lambda_k)), m = j, k\} \tag{67}$$

and

$$l = \begin{cases} 1, & \text{for } j = 1, k = 2, \\ 2, & \text{for } j = 2, k = 3, \\ 3, & \text{for } j = 1, k = 3. \end{cases} \tag{68}$$

4. Calculate the intersection, (α_1^I, α_2^I) of $\rho_{\lambda_1}, \rho_{\lambda_2}, \rho_{\lambda_3}$ if it exists. If $\rho_\lambda(\alpha_1^I, \alpha_2^I) < 1$, form the union of (α_1^I, α_2^I) with OP and the union of $\rho(\alpha_1^I, \alpha_2^I)$ with OQ.

5. Evaluate $\rho_{\lambda_m}(\alpha_1^*(\lambda_j, \lambda_k), \alpha_2^*(\lambda_j, \lambda_k))$ for $\lambda_m \neq (\lambda_j \text{ or } \lambda_k)$ for all (j, k) pairs in Eq. (68) (i.e., for all values of l). If $\rho_{\lambda_m} > \rho_l^*$ for the lth entry in OQ then delete ρ_l^* from OQ and the lth entry, $(\alpha_1^*(\lambda_j, \lambda_k), \alpha_2^*(\lambda_j, \lambda_k))$, of OP.

Now OP contains all possible optimal vertices and minima of F_M and OQ contains the corresponding value of F_M at each entry of OP.

6. Determine $(\alpha_1^{**}, \alpha_2^{**})$ as the pair in OP which corresponds to the entry in OQ which is the minimum over all members of OQ (i.e., if the lth entry in OQ is the minimum in OQ then $(\alpha_1^{**}, \alpha_2^{**})$ is the lth entry of OP).

By taking $N = 3$ we have made the brute force optimization procedure affordable and, as we mentioned, increasing N does not necessarily yield any benefit.

9. Application of Optimal Back-Average Algorithm to the Outer Loop of the AD Model

The two quantities back-averaged in the outer loop of the AD model are $V_{\text{sep}}(t)'$, which establishes the location of the separatrix in volume space, and J, which is an input for the equilibrium solver. One way of applying the optimal back-average algorithm (OBA) to the outer loop simply uses the same optimal back-average parameter α_{opt} for $V_{\text{sept}}(t)'$ and J. This method works well for easy problems.

A somewhat more complicated method of back-averaging the outer loop is to use OBA to calculate a back-average parameter, α_{sep}, for $V_{\text{sep}}(t)'$ and, then, to calculate a different one, α_J, for back-averaging J. This use of two separate back-average parameters is essential for convergence in the hard cases (small resistivity).

First, take the same α to back-average both $V_{\text{sep}}(t)'$ and J. We begin by solving for A^{-1}, where A is the column matrix,

$$A = [\bar{X}_{n-3}\bar{X}_{n-2}\bar{X}_{n-1}], \tag{69}$$

where

$$\bar{X}_n = \begin{vmatrix} V_{\text{sep}}^{(n)}(t)' \\ K_{\text{in}}^{(n)} \\ K_{\text{out}}^{(n)} \end{vmatrix}; \tag{70}$$

here $V_{\text{sep}}^{(n)}(t)'$ is $V_{\text{sep}}(t)'$ at the nth iteration; $K_{\text{in}}^{(n)} = K(V_{\text{sep}}^{(n)}/2)$ and $K_{\text{out}}^{(n)} = K((V_{\text{plasma}}^{(n)} - V_{\text{sep}}^{(n)})/2 + V_{\text{sep}}^{(n)})$ at the nth iteration of the outer loop. We form

$$\tilde{T} = BA^{-1}, \tag{71}$$

where

$$B = [\bar{X}_{n-2}\bar{X}_{n-1}\bar{X}_n]. \tag{72}$$

Next, the eigenvalues of $\tilde{T}$ are calculated. Then (24) is used to get $\lambda_i^{(n)}$. The eigenvalues $\lambda_i^{(n)}$, $i = 2, 3$ are the arguments of OBA. When SBA is used no additional iterations are required to calculate the eigenvalues. However, when 3PBA is applied, three additional iterations of SBA with a fixed α must precede the calculation of the $\lambda_i^{(n)}$ at the nth iteration.

Two of the three eigenvalues of $\tilde{T}$, $\tilde{\lambda}_2^{(n)}$ and $\tilde{\lambda}_3^{(n)}$, are used by OBA. The third eigenvalue $\tilde{\lambda}_1^{(n)}$ approximates 1. This is expected since for a fixed point of (23) to exist there must be an eigenvalue of $\tilde{T}$ near 1. The third eigenvalue used as input to OBA is $\lambda_0 = 0$. This in included to ensure that eigenvalues of T near zero which would have converged with straight iteration remain convergent. This follows from continuous dependence.

In brief, the outer loop iteration with OBA is:

1. Back-average $V_{sep}(t)'$ and J with α_{start}, a starting back-average parameter, for three iterations.

2. Calculate $\tilde{T}$, $\tilde{\lambda}_i^{(n)}$, $i=2, 3$ and then $\lambda_i^{(n)}$, $i=2, 3$.

3. Input to OBA $\lambda_1=0$, $\lambda_2^{(n)}$ and $\lambda_3^{(n)}$ and specify SBA or 3PBA to get α or (α_1, α_2), respectively.

4 Do Case A or B:

Case A. SBA, use α for three iterations then go to step 2.

Case B. 3PBA, use (α_1, α_2) for the next $2k$ iterations, $k=1, 2, 3$. Iterate with α fixed, $\alpha=\alpha_{safe}$, for 3 iterations. Go to step 2.

To demonstrate the practical benefits of using OBA in this way in the outer loop, we compare ratios of convergence for cases of increasing difficulty. For this comparison we define the convergence factor, CF, of the iteration scheme over iterations i through k as

$$CF = \exp(\ln(e_k/e_i)/(k-i)), \tag{73}$$

where e_k represents $\|J_k - J_{k-1}\|$ at the kth iteration of the outer loop.

We now consider an "easy" case, one which converges, perhaps very slowly, without back-averaging. The resistivity for the easy case is two orders of magnitude larger than that for the "hard" cases. The J profile in the solution of the easy case is not sharply peaked near the separatrix. Moreover, the modulation of the shaping coils is decreased by two orders of magnitude from the hard cases. The excursion and velocity of the separatrix are small enough so that straight iteration of the outer loop converges.

Convergence factors for the easy problem resulting from using various back-average schemes are shown in Table I. The overall convergence is calculated with $i=1$ and k equal to the total number of iterations required to meet the convergence criteria. The tail convergence, CF_{tail}, is calculated with $i=k-3$.

The overall convergence factors in Table I clearly show that back-averaging is necessary for practical convergence since the convergence factor of the iteration without back-averaging (0.9979) is so close to 1 that 5800 iterations are required for convergence to a final error of 1×10^{-6} from a typical initial error of 2×10^{-1}. The Marder–Weitzner iteration scheme is significantly better than no back-averaging. However, for this problem it, too, is slow since 750 iterations are necessary to meet the convergence criterion with a CF of 0.984.

Simple back-averaging with any constant back-average parameter, α in the range from 0.05 to 0.5 will result in reasonable convergence. Better, the optimum constant back-average parameter of 0.25 yields fast convergence: it requires only thirteen iterations. The cost of optimizing α for a time step is iteration of that step to convergence with two or three different back-average parameters. This optimum

TABLE I

CF for an Easy Case

Case	Iteration Method	$CF_{overall}$	CF_{tail}
1	Straight iteration $\alpha = 0$	0 998	0 998
2	Marder–Weitzner iteration	0 984	0 985
3	OBA simple back-average, SBA	0 330	0 200
4	OBA three-point back-average, 3PBA	0 324	0 206
5	Constant simple back-average $\alpha = 0\,05$	0 760	0 783
6	Constant simple back-average $\alpha = 0\,15$	0 412	0 315
7	Constant simple back-average $\alpha = 0\,25$	0 363	0 321
8	Constant simple back-average $\alpha = 0\,5$	0 575	0 487

Note We show $CF_{overall}$, the convergence factor for the entire outer loop from first estimate to convergence and the convergence factor CF_{tail}, for the last three iterations before convergence All CF are for the same time step of an easy case The shaping coil modulation was 0 0002, the resistivity, 2

changes from one time step to another for the same resistivity and modulation and changes even more for runs with different resistivity and modulation.

The benefit of both the OBA with simple back-average and the OBA with three-point back average schemes is clear. Their CF's are at least 12% smaller than any others yielding convergence in at most eleven iterations. Most of all, the automatic routines *do not* need two or more preliminary calculations of a single time step to find the optimum constant back-average parameter, after which an additional calculation would have to be done to take advantage of the optimum constant back-average parameter. Instead, both OBA routines use information obtained from previous iterates of the same time step to adjust α for the following iteration of that time step. Thus, each time step is solved only once, not two or more times. Accordingly, in this comparison both OBA routines are more than two times faster.

The above comparison is striking but, in practice, somewhat unrealistic because

third time. One would solve the time step only once with an average value of α. When this average case comparison is made, the automatic back-average routines yield a CF approximately 50% smaller than the average CF of all the constant back-average parameter cases. Comparing the average number of iterations to convergence, we note that with constant α, nineteen iterations are required, while only eleven are required for the automatic routines. Accordingly, on the average, OBA is almost twice as fast as using a constant back-average routine. For both OBA routines, CF_{tail} is markedly smaller than $CF_{overall}$. This is expected because the outer loop transformation is better approximated by a linear transformation near the fixed point. Once in a neigborhood of a solution we can obtain high accuracy at minimal cost.

Now let us back-average the outer loop using OBA for V_{sep} and J, independently. Let α_{sep} be the back-average parameter for SBA applied to $V_{sep}^{(n)}(t)'$. To determine

α_{sep} we use (43) with λ^n given in terms of $\tilde{\lambda}^{(n)}$ by (24) with α fixed in the nth and $(n-1)$th iteration. In this,

$$\tilde{\lambda}^{(n)} = \frac{\Delta_n}{\Delta_{n-1}}, \tag{74}$$

where $\Delta_n = V_{sep}^{(n)}(t)' - V_{sep}^{(n-1)}(t)'$. The method is equivalent to using only one eigenvalue as input to OBA and specifying SBA as the back-average method. When α is not constant for the two successive iteration steps, $\alpha_{sep}^{(n)} \neq \alpha_{sep}^{(n-1)}$, Eq. (74) is not applicable. In that event, let $V_J' = V_{sep}^{(n+3-J)}(t)'$ and $\alpha_J = \alpha_{sep}^{(n+3-J)}$. Use (23) to iterate twice by

$$\tilde{T}(V_J' - V_0') = \lambda^{(n)}(V_J' - V_0') = (V_{J+1}' - V_0'), \qquad J = 1, 2,$$

and eliminate V_0' to find

$$\lambda^{(n)} = \frac{(1-\alpha_2)[V_3' - \alpha_3 V_2'] - (1-\alpha_3)[V_2' - \alpha_2 V_1']}{(1-\alpha_3)(1-\alpha_2)(V_2' - V_1')}. \tag{75}$$

Next, OBA is used on J. We calculate either α_J or the pair $(\alpha_1, \alpha_2)_J$ depending on whether SBA or 3PBA is to applied to the J profile of the outer loop. The eigenvalues used by OBA to calculate α_J or $(\alpha_1, \alpha_2)_J$ are calculated from (69), (71), and (72), where

$$\bar{X}_n = \begin{vmatrix} \psi_{sep}^{(n)} \\ K_{in}^{(n)} \\ K_{out}^{(n)} \end{vmatrix} \tag{76}$$

and $\psi_{sep}^{(n)}$ is ψ at the separatrix at the nth iteration.

The choice of variables in $\bar{X}_n$ depends on the problem being solved. In our solutions of the AD cases, the above variables work well but other combinations of ψ, K, and V also work. As a guideline to choosing $\bar{X}$ we ask that one or two variables should account for convergence over a major part of the domain and the third should control the convergence for any localized exceptional feature. Here, in (76), K is associated with convergence of quantities which are *not* near the separatrix and ψ_{sep}, near the separatrix. Certainly, some experience and intuition about the problem are required to achieve convergence of a system with over 8000 variables with a representative subsystem of only 3. Note that we cannot use a Newton method for our problem, since we know of no way of defining a derivative for the outer loop. A secant method is also impractical because it requires solving for the Jacobian of a system with 8000 variables, more work than solving the entire problem.

The motivation for choosing the α_J and α_{sep} independently is twofold. First, the geometry is extremely sensitive to $V_{sep}(t)'$ in typical hard cases. If special care is not taken to control $V_{sep}(t)'$, the geometry and profiles will quickly diverge from the

equilibrium solution of the previous time step and the outer loop will never converge. Second, after about five iterations of the geometry and profiles, the shape begins to converge much faster than the profiles. Since $V_{sep}^{(n)}(t)'$ is more closely related to the shape than the J profile, it too will converge faster than the profiles. This difference in the rate of convergence can be converted from a hindrance to an advantage if $V_{sep}^{(n)}(t)'$ is back-averaged separately from J. Once the shape and $V_{sep}(t)'$ have converged to a moderate degree of accuracy, further minor variations in them need not be considered in the choice of α_J. Instead, α_J can be determined by the flux profile which dominates the end of the iteration.

The AD formulation of the model makes it simple to treat $V_{sep}^{(n)}(t)'$ and J separately. This is one of the major advantages of introducing volume space normalized to $V_{sep}^{(n)}(t)$ in the diffusion equation to locate the separatrix. (Recall that $V_{sep}^{(n)}(t)$ is interpolated as a continuous cubic function with continuous derivative $V_{sep}(t)'$.)

The above method complicates the outer loop somewhat because we can change only one back-average parameter at a time. Either α_{sep} must be held fixed while α_J is calculated or vice versa; otherwise it is not clear to which change the eigenvalues are reacting. Accordingly, the outer loop is:

1. Set $V_{sep}^{(n)}(t)' = V_{sep}(t-dt)'$ and iterate the outer loop using $\alpha_J = \alpha_{1\,start}$ to back-average J for a few iterations, say, $k_1 \leqslant 8$.

2. Iterate the outer loop with $\alpha_{sep} = a_{2\,start}$, α_J unchanged for $k_2 = 3, 4, 5$ or until the error, **er**, in the outer loop, is less than $e_{1\,start}$.

3. Calculate $\lambda_{sep}^{(n)}$ then α_{sep} from OBA and back-average for $k_2 = 2, 3, 4$ or 5 iterations. Repeat this step until **er** $< e_{2\,start}$.

4 Fix α_{sep} to its last value or to $\alpha_{1\,start}$ (By this time $V_{sep}(t)'$ has converged to the point where α_{sep} is no longer critical.)

5. If **er** $<$ **erc** then exist (**erc** is our fixed error tolerance).

6. Calculate $\lambda_J^{(n)}$ then α_J or $(\alpha_1, \alpha_2)_J$ from OBA.

7. Do Case A or Case B

Case A. SBA, use α_J for 3 iterations. Go to step 5.

Case B. 3PBA, use $(\alpha_1, \alpha_2)_J$ for $2k_3$ iteration, $k_3 = 1, 2, 3$. Bach-average the outer loop with $\alpha_J = \alpha_{J\,safe}$, α_{sep} unchanged. Go to step 5.

The choice of $\alpha_{1\,start}$, $\alpha_{2\,start}$, $e_{1\,start}$, $e_{2\,start}$, k_1, k_2, and k_3 depends on the case being solved. Typically the closer, $\alpha_{1\,start}$ and $\alpha_{2\,start}$ are to 1, the safer they are for difficult cases and the slower the convergence, even for less difficult cases. We find that $\alpha_{1\,start} = 0.5$ works well for most cases. $\alpha_{2\,start} = 0.9$ works for all cases but $\alpha_{2\,start} = 0.7$ works better for less difficult cases; however, $\alpha_{2\,start} < 0.9$ does not converge at all for difficult cases (see Table III). The choice of $e_{1\,start}$ is not critical; we favor $e_{1\,start} = 5 \times 10^{-2}$. Good values for $e_{2\,start}$ are approximately 1×10^{-3}. If $e_{2\,start}$ is too big, the eigenvalues input to OBA have little meaning since the outer loop transformation is poorly approximated by the linear transformation $\tilde{T}$ and the

linear assumption upon which OBA is based is not applicable. Experience shows $K_1 = 5$, $K_2 = 4$, and $K_3 = 1$ or 2 work well in all cases. $\alpha_{J\text{safe}} = 0.6$ or 0.7 work well, in all but the difficult cases, where $\alpha_{J\text{safe}} = 0.8$ is advised.

In addition to the above parameters, we specify three other parameters, $\alpha_{\min}$, $\alpha_{\max}$, and $r_{\max}$. We admit only back-average parameters greater than $\alpha_{\min}$; typically $\alpha_{\min} \sim -0.5$. The reason $\alpha_{\min}$ is specified is that only three eigenvalues are input into OBA but, in fact, T has many more. Back-averaging with $\alpha < 0$ is beneficial for $0 < |\lambda| < 1$ but can cause eigenvalues with larger absolute value to diverge. $\alpha_{\max}$ is the maximum α permitted; usually $\alpha_{\max}$ lies between 1.2 and 1.5. $r_{\max}$ is used to determine whether the maximum $\rho(\lambda_i^{(n)})$, $i = 0, 2, 3$, for the α_J output from OBA is acceptable. If $\rho(\lambda) > r_{\max}$, then $\alpha_{J\text{safe}}$ is used instead of α_J.

A comparison of convergence factors for a typical adiabatic limit case is presented in Table II. A comparison of Methods 1 and 6 shows that the use of OBA to obtain α_{sep} and $(\alpha_1, \alpha_2)_J$ cuts the total number of iterations to convergence by

TABLE II

Convergence Factor Comparison for a Typical Adiabatic Limit Case

Method	Back-average method	CF	CF_{tail}	Total iterations	Iteration up to step 4
1	$\alpha_{\text{sep}} = 0.5$ $\alpha_J = 0.9$ fixed	0.874	0.888	96	—
2	$\alpha_{\text{sep}} = 0.5$ $\alpha_J = 0.8$ fixed	0.812	0.791	61	—
3	$\alpha_{\text{sep}} = 0.5$ $\alpha_J = 0.7$ fixed	0.820	0.824	65	—
4	$\alpha_{\text{sep}} = 0.5$ α_J = MWS (Not calculated to convergence)	0.998	0.998	6443	—
5	α_{sep} OBA α_J OBA SBA $\alpha_{\text{start}} = 0.9$	0.838	0.7789 (0.785)[a]	56	30
6	α_{sep} OBA $(\alpha_1, \alpha_2$ OBA, 3PBA $\alpha_{\text{start}} = 0.9$	0.762	0.415 (0.666)[a]	45	30
7	α_{sep} OBA $(\alpha_2, \alpha_2)_J$ OBA, 3PBA $\alpha_{\text{start}} = 0.7$	0.701	0.478 (0.640)[a]	35	21

[a] 0.779 (0.785) indicates $CF_{tail} = 0.779$ and $CF_{OBA} = 0.785$, where CF_{OBA} is calculated with i equal the number of iteration to $e_{2\text{start}}$.

Note Contains the convergence factors for the entire outer loop, CF as well as CF_{tail} for a typical near-adiabatic case The total number of iterations and the number of iterations of step 2 of the outer loop using OBA are enclosed where appropriate The same time step is done for all cases with $\eta = 0.025$ and the shaping coil modulation equal to 0.02 (Iteration counts are calculated assuming an initial error of 2×10^{-1} and a final error of 1×10^{-6} for the OBA cases $e_{2\text{start}} = 5. \times 10^{-4}$).

more than one-half. Similarly the number of iterations for Method 7 using OBA is only slightly more than one-half of the number of iterations of Method 3 which uses a fixed α_{sep} and α_J equal to the $\alpha_{1\,start}$ and $\alpha_{2\,start}$, respectively, of Method 7. Furthermore, the slowest case with 3PBA and OBA, Method 6, converged in two-thirds the number of iterations required for Method 2, the fastest routine without OBA.

This method of calculating α_{sep} and α_J or $(\alpha_1, \alpha_2)_J$ separately was used in the "easy case" which was discussed earlier. The resulting convergence factor for OBA with 3PBA was 0.321. The equivalent procedure using OBA with the same α for α_{sep} and α_J resulted in an almost identical CF, 0.324. (See Method 4 of Table I.) For the method using OBA with SBA to calculate α_{sep} and α_J separately, CF = 0.377; the equivalent procedure using the same α_{sep} and α_J is only slightly smaller with CF = 0.330. (See Method 3 of Table I.)

Notice that a smaller $\alpha_{J\,start}$ can reduce the startup number of iterations as shown by Method 7, $\alpha_{J\,start} = 0.7$. However, more difficult cases occur at different time steps of the same run, and for these, $\alpha_{J\,start} < 0.9$ never converges (See Table III). Accordingly, to be safe when using the same $\alpha_{J\,start}$ for an entire run, typically three cycles of the forcing oscillations, a larger $\alpha_{J\,start}$ is required while a smaller $\alpha_{J\,start}$ can be used to accelerate the convergence of some time steps. In practice, we use $\alpha_{J\,start} = 0.7$ when it is possible to watch the convergence periodically and change $\alpha_{J\,start}$ if necessary. We use $\alpha_{J\,start} = 0.9$ for the overnight runs which are done completely hands off.

Although the convergence in the preliminary iteration is somewhat slow, the use of OBA after $\mathbf{er} < e_{2\,start}$ enables rapid convergence to within a small tolerance. The tolerance in this case is 1×10^{-6}. Notice that if the tolerance is reduced by 2 orders of magnitude to $1. \times 10^{-8}$ only eleven more iterations are required for convergence based on $CE_{OBA} = 0.67$ for Method 6, in contrast to the 36 iterations required if OBA is not used as in Method 1. This saves two-thirds the cost of the additional accuracy.

This case and similar ones were tried with only one back-average parameter or pair for both $V_{sep}^{(n)}(t)'$ and $J^{(n)}$ but divergence and convergence were so slow and erratic that they were unpredictable. This is due to the difference in the rate of convergence of the shape from the rate of convergence of the profiles. To fine tune $V_{sep}^{(n)}(t)'$ first, then $J^{(n)}$, works well and enables us to use the AD code for cases with singular profiles and rapidly changing shapes, problems which hitherto were impossible to solve.

Table III. This difficult case is one time step which requires over 150 iterations of the outer loop with optimum fixed back-average parameters. We present the case $\eta = 0.025$ and fix the coil modulation, $F_{mod} = 0.02$. At this time step, $V_{sep}(t)'$ is near its maximum. The optimum fixed back-average parameters for this time step are $\alpha_{sep} = 0.5$ and $\alpha_J = 0.9$. To go from an initial error of 2×10^{-1} to a final error of 1×10^{-6} required 152 iterations of the outer loop with optimum fixed back-average parameters. Savings of 50% result from using OBA to calculate α_{sep} and to calculate either α_J for SBA or the pair $(\alpha_1, \alpha_2)_J$ for 3PBA in the outer loop of the

TABLE III

Convergence Factor Comparison for a Difficult Adiabatic Limit Case

Method	Back-average method	CF	CF_{tail}	Total iteration	Iteration up to step 4
1	$_{sep} = 0.5$ $\alpha_J = 0.9$ fixed	0.923	0.960	152	—
2	$\alpha_{sep} = 0.5$ $\alpha_J = 0.875$ fixed	1.010	—	—	—
3	$\alpha_{sep} = 0.5$ $\alpha_J = 0.8$ fixed	1.011	—	—	—
4	$\alpha_{sep} = 0.5$ $\alpha_J = 0.7$ fixed	1.012	—	—	—
5	$\alpha_{sep} = 0.5$ $\alpha_J =$ MWS (Not calculated to convergence)	0.998	0.998	6100	—
6	α_{sep} OBA, SBA $(\alpha_1, \alpha_2)_J$ OBA, 3PBA $\alpha_{J\,start} = 0.9$	0.852	0.816 (0.818)[a]	76	45
7	α_{sep} OBA, SBA α_J OBA, 3PBA $\alpha_{J\,start} = 0.9$	0.850	0.822 (0.813)[a]	75	45

[a] 0.816 (0.818) indicates $CF_{tail} = 0.816$, $CF_{OBA} = 0.818$

Note Contains the convergence factors, CF, and the total number of iterations for the entire outer loop from initial error of $2 = 10^{-1}$ to convergence of 1×10^{-6} for a difficult case The tail convergence factor, CF_{tail} as well as CF_{OBA} (see Table II for notation) are presented The same time step is done for all cases The shaping coil modulation was 0.02 The resisistivity was 0.025. The time step is chosen for $V_{sep}(t)'$ near its maximum

same time step. Notice that convergence for MWS is so slow that it is not usable. With the same parameter set ($\alpha_{1\,start}$, α_{2start}, $e_{1\,start}$, e_{2start}, α_{max}, α_{min}, r_{max}) for both low resistivity and the extreme case of low resistivity combined with high modulation, using OBA results in a savings of at least 50% over the fixed α cases. Therefore, it is cost effective to run our AD code completely hands off. Even when compared to a hands-on method which uses the optimum fixed α_{sep} and α_J for each time step (see Method 2 of Table II and Method 1 of Table III) the hands-off method using OBA shows a savings of over 40%. (see Method 6 in both Tables II and III.) Almost twice as many runs can be made using OBA to calculate α_{sep} and $(\alpha_1, \alpha_2)_J$ than with optimum *fixed* α_{sep} and α_J. In addition, the outer loop using OBA can be done hands off.

SUMMARY

We have analyzed two- and three-point back-averaging as multi-stage iterative techniques for finding a fixed point of a nonlinear function F. For the complex

eigenvalues of the linear approximation to F, we calculated the domain of convergence in the complex λ-plane as a function of the back-average parameter in the neighborhood of the fixed point. For two-point back-averaging convergence is always possible provided only that all eigenvalues fall to one side of the line $\mathrm{Re}\,\lambda = 1$. The Marder–Weitzner scheme (the first we know to converge with eigenvalues on both sides of one) is now seen to be a special case of the more general method, three-point back-averaging (3PBA) An analytic formulation gives the domain of convergence of the method as a function of the back-average parameter; the domain boundaries are members of a one-parameter family of curves containing lemniscates and ovals of Cassini.

We have found an optimal back-average algorithm (OBA) which monitors the dominant eigenvalues of F (assuming F is approximately linear) during successive iterations. Our techniques are capable of solving a wide variety of problems of current interest in fusion plasma dynamics. In particular, we have employed the alternating dimension algorithm to solve resistive MHD equations of motion representing the time evolution of complicated magnetic field flux surfaces. The back-averaging technique provides a fast and accurate solution. The central difficulties appear in the so-called outer loop of the algorithm, which calculates the strongly peaked current density profiles and determines the position and velocity of a singular flux surface or separatrix We considered the outer loop iteration a three-dimensional transformation, F, to approximate the eigenvalues of its linearization and solved the evolution problem affected by these eigenvalues, using optimized back-average parameters.

Convergence tests show that for practical convergence of the AD outer loop in a doublet topology back-averaging is essential. Moreover, the tests of OBA show a double rate of convergence obtained with a fixed optimal back-average parameter. This saving is further enhanced since each time step is solved only once using OBA whereas finding the fixed optimal back-average parameter requires solving one time step at least twice and perhaps three times. Most important, once we arrive in the neighborhood of a solution, OBA yields accurate convergence with a minimum of additional work.

ACKNOWLEDGMENTS

This work is dedicated to the memory of Dr Harold Grad He provided the framework of the material contained herein, and through his genius supplied the impetus necessary to complete this work To Dr Albert Blank I give my profound thanks and gratitude for the countless hours he spent offering both technical and editorial assistance and advice His expertise was crucial In addition, I am deeply grateful to Dr William Grossmann for his support and advice. Many thanks also to Dr Donald Stevens and Dr Pung Nien Hu for their discussions concerning the numerical and analytic formulation of the alternating dimension algorithm. And to Dr Kurt Riedel, Dr Juergen Gross, and Dr Marsha Berger my sincerest appreciation for their help in formulating the back-average analysis This work was supported by U S Department of Energy Grant DE-FG02-86ER53223

REFERENCES

1 H GRAD AND J HOGAN, *Phys Rev Lett* **24**, 1337 (1970)
2 Y-P PAO, *Phys Fluids* **19**, 1177 (1967)
3 S P HIRSHMAN AND S C JARDIN, *Phys Fluids* **22**, 731 (1979)
4 H GRAD, "Classical plasma diffusion," Conference on Plasma Theory, Kiev, October 1971, Courant Institute of Mathematical Sciences Report MF-69, New York University (unpublished)
5 H GRAD, P N HU, AND D C STEVENS, *Proc Natl Acad Sci USA* **72**, 3789 (1975)
6. H GRAD, P N HU, D STEVENS, AND E TURKEL, *Plasma Phys Controlled Nucl Fusion Res* **2**, 355 (1976).
7 H GRAD AND P N HU, "Classical diffusion Theory and simulation codes," Courant Institute of Mathematical Sciences, Report MF-91 New York University, (1978) (unpublished)
8 J MOSSINO AND R TEMAM, *Duke Math J* **43**, 475 (1981)
9 P LAURENCE AND E W STRENDULINSKY, *Comm Pure Appl Math* **38**, 333 (1985)
10 G BATEMAN, *Nucl Fusion* **13**, 227 (1973)
11 H GRAD, "Survey of $1\frac{1}{2}$-D transport codes," Courant Institute of Mathematical Sciences Report MF-93, New York University, 1978 (unpublished)
12 J BLUM AND J LE FOLL, *Comput Phys Rep* **1**, 465 (1983)
13 T H JENSEN, A M SLEEPER, AND J L LUXON, Excitation and investigation of low frequency modes in Doublet III, *Bull Amer Phys. Soc* **25**, No 8 (1980)
14 H R SCHWARZ, H RUTISHAUSER, AND E STIEFEL, *Numerical Analysis of Symmetric Matrices* (Prentice-Hall, Englewood Cliffs, NJ, 1973), Chap 2
15 W NIETHAMMER AND R S VARGA, *Numer Math* **41**, 177 (1983)
16 B MARDER AND H WEITZNER, *Plasma Phys* **12**, 435 (1970)
17 W MURRAY AND M OVERTON, *SIAM J Sci Stat Comput* **1**, 345 (1980)