Mathematical Modelling, Vol. 5, pp. 275–297, 1984
Printed in the USA.

0270-0255/84 $03.00 + 0

BLOCK DIAGONAL DOMINANCE FOR SYSTEMS OF NONLINEAR EQUATIONS WITH APPLICATION TO LOAD FLOW CALCULATIONS IN POWER SYSTEMS

MARIJA ILIĆ-SPONG*

Department of Electrical and Computer Engineering
University of Illinois at Urbana
Urbana, Illinois 61801, USA

I. NORMAN KATZ

Department of Systems Science and Mathematics
Washington University
St. Louis, Missouri 63130, USA

HUIZHU DAI

The Northeast Electric Power Institute
Jilan, People's Republic of China

JOHN ZABORSZKY

Department of Systems Science and Mathematics
Washington University
St. Louis, Missouri 63130, USA

Communicated by Ervin Y. Rodin

(*Received November* 1983)

Abstract—The concept of a pointwise strict (or Ω) diagonally dominant nonlinear function, first introduced by Moré, is generalized to the blockwise case. A sufficient condition is obtained for the convergence of underrelaxed block Jacobi and block Gauss-Seidel iterations for a nonlinear system of equations in terms of the strict (or Ω) diagonal dominance of an associated matrix. A new formulation for the determination of the steady-state load flow in lossless electric power systems is described and it is shown that this formulation leads to the solution of a system of quadratic equations in the unknown (complex-valued) voltages. Under suitable assumptions on the power system the sufficient condition is satisfied. Numerical examples, consisting of an illustrative three bus system and a realistic thirty bus system, are presented. Results of our block Gauss-Seidel iteration are compared with those of Newton-Raphson iteration.

Keywords: iterative solution, Jacobi and Gauss-Seidel iterations, nonlinear systems of equations, load flow in electric power systems, block methods.

This research was supported by the Department of Energy under contract numbers DEAC 01-78ET-29636 and DEAC 01-79ET-29637.

* The work of this author is an extension of part of her doctoral dissertation at the Department of Systems Science and Mathematics, Washington University, St. Louis Missouri 63130, USA.

1. INTRODUCTION

The notions of strict diagonal dominance and irreducible diagonal dominance have long been used in solving linear systems of equations via the iterative methods of Jacobi and Gauss-Seidel. In the work of Moré [7] the notion of diagonal dominance of matrices is generalized to systems of nonlinear equations. A (pointwise) diagonally dominant nonlinear mapping is defined for the real domain, and it is shown that this definition is equivalent to diagonal dominance of matrices when the mapping is linear. The notion of irreducible diagonal dominance is also generalized to Ω diagonal dominance. Sufficient conditions are given for the convergence of pointwise nonlinear Jacobi or Gauss-Seidel methods.

In this paper we extend the definition of a *pointwise* strictly (or Ω) diagonally dominant nonlinear mapping to a *blockwise* strictly (or Ω) diagonally dominant nonlinear mapping. We then derive sufficient conditions for a *quadratic* mapping to be a blockwise strictly (or Ω) diagonally dominant mapping in terms of an associated blockwise strictly (or Ω) diagonally dominant matrix. A sufficient condition is also derived for more general nonlinear mappings, but in the nonquadratic case this condition is more difficult to verify in practice.

Feingold and Varga [3] have studied block diagonal dominance for linear mappings and their work has been applied in practice (for example, by Price in [8]). The condition that we derive here is similar to that in [3] but it applies to quadratic mappings as well as linear.

Our motivation for introducing blockwise diagonally dominant nonlinear mappings comes from load flow studies in electric power systems. Every electric power system consists of a set of hydro, thermal, and/or nuclear generating units, various types of loads, and a transmission system connecting the generating units and the loads. For such a system the desired performance is one in which, for known loads and available generating units, the power balance at each instant is satisfied without violating any of the given constraints on produced and transmitted power.

Given the steady state value of the injections (outputs of generators, loads, etc.) at the buses of a symmetrical three phase electric power system, the load flow problem is that of determining the magnitudes and phase angles of the voltages at the buses. This problem has a vast literature (see, for example, [1, 4, 11]) and is of considerable importance in practice. The relations between the voltage and current in the transmission network are mostly linear (strictly linear, if only transmission lines are involved [9]); however, power relations on the network are inherently nonlinear and some of the components of injections, loads for example, are nonlinear functions of the bus voltages. The load flow problem is then the solution of a large number of nonlinear algebraic equations.

In [6] a new formulation of the load flow problem is presented. It is shown there that solving the load flow problem is equivalent to solving a system of *quadratic, nonanalytic* equations in $\mathscr{C}^n$. (The nonanalyticity occurs because of the presence of terms of the for $|\hat{E}|^2$, where $\hat{E}$ is complex, $\hat{E} = E_x + \sqrt{-1}\, E_y$). Because of the nonanalyticity, the complex derivative does not exist, so that standard iterative procedures which use complex derivatives do not apply. Instead we introduce (2×2) blocks on R^{2n} via block Gauss-Seidel or block Jacobi methods.

One could use the sufficient pointwise conditions given in [7], applied to quadratic mappings, in order to obtain $2n$ conditions for convergence. However, since each node in the system has associated with it two real valued unknown voltages, it is more natural to consider the nonlinear mapping from $\mathscr{C}^n$ into $\mathscr{C}^n$ as consisting of (2×2) blocks in R^{2n} and to introduce the blockwise diagonal dominance conditions. This results in n conditions instead of $2n$. Moreover, since our sufficient conditions in this case involve only norms of 2×2 matrices they are very easy to verify.

In Sec. 2 of this paper we define a blockwise strictly (or Ω) diagonally dominant function F and we derive a sufficient condition for F to be a blockwise strictly (or Ω) diagonally dominant function in terms of strict (or Ω) diagonal dominance of an associated matrix. In Sec. 3, this condition is applied to obtain sufficient conditions for the convergence of block Jacobi or Gauss-Seidel iterations. This application is straightforward once the results in Sec. 2 have been established and follows the reasoning in Moré [7]. In Sec. 4 we give a concise description of the $\hat{S}$–$\hat{E}$ model of the load flow problem for electric power systems. This model is new and contains several novel features [6]. In Sec. 5, we show how the notion of Ω block diagonal dominance is applied to solve the system of quadratic equations which arise in the $\hat{S}$–$\hat{E}$ model. Part of the application depends upon certain physical assumptions, and we state clearly what these assumptions are. Finally, we conclude in Sec. 6 with numerical examples consisting of an illustrative, small, three-bus electric power system, as well as a realistic 30 node system used as a standard IEEE test by AEP (American Electric Power). Comparisons are made with the conventional Newton-Raphson iteration.

2. DEFINITIONS AND BASIC RESULTS

We first review the motivation for the definition of a pointwise diagonally dominant function [7]. For $\mathbf{x} \in R^n$, let $\| \mathbf{x} \|_\infty = \max_i | x_i |$.

Lemma 1. Let $\mathbf{v} \in R^n$; then the following statements are *equivalent*:

(a) $| v_k | > \sum_{j \neq k} | v_j |$ for some $k \in N$, $N = \{1, 2, 3 \ldots n\}$.

(b) $\sum_{j=1}^{n} v_j x_j = 0$ implies that $| x_k | < \| \mathbf{x} \|_\infty$ for any $\mathbf{x} \in R^n$.

For a proof see [7]. If A is an $n \times n$ matrix, and for some $k \in N$, $v_j = a_{kj}$ $j = 1 \ldots n$, then (a) is equivalent to assuming "strict diagonal dominance on the kth row". Condition (b), however, can be generalized to the nonlinear case.

Definition 1

(a) A functional f_k: $D \subset R^n \rightarrow R^1$ is pointwise (strictly) diagonally dominant on D with respect to the kth variable if for every $\mathbf{x} \neq \mathbf{y}$ in D

$$f_k(\mathbf{x}) = f_k(\mathbf{y}) \text{ implies that } | x_k - y_k | \leq \| \mathbf{x} - \mathbf{y} \|_\infty \ (<)$$

(b) A *function* F: $D \subset R^n \rightarrow R^n$ is pointwise (strictly) diagonally dominant on D if for each $k \in N$ the kth component function of F, f_k, is pointwise (strictly) diagonally dominant with respect to the kth variable.

In order to define a *blockwise* strictly diagonally dominant function we first partition a vector $\mathbf{x} \in R^n$ into subvectors. Let

$$\mathbf{x}^T = (\mathbf{x}_1^T, \mathbf{x}_2^T \ldots \mathbf{x}_s^T), \qquad \mathbf{x}_i \in R^{n_i}, \quad \mathbf{x}_i^T = (x_{i1} \ldots x_{in_i}) \tag{1}$$

$$\sum_{i=1}^{s} n_i = n$$

and let

$$\| \mathbf{x} \|_\infty \equiv \max_{i \in S} \| \mathbf{x}_i \|_\infty = \max_{i,j \in N} | x_{ij} | \qquad S = \{1, 2 \ldots s\} \tag{2}$$

$$\| \mathbf{x} \|_1 \equiv \sum_{i=1}^{s} \| \mathbf{x}_j \|_1 = \sum_{i=1}^{s} \sum_{j=1}^{n_i} | x_{ij} | = \sum_{i,j \in N} | x_{ij} | \tag{3}$$

Let A be an $n \times n$ matrix partitioned into $n_i \times n_j$ matrices A_{ij}, $1 \leq i, j \leq s$:

$$A = \begin{bmatrix} A_{11} & A_{12} & \ldots & A_{1s} \\ A_{21} & A_{22} & \ldots & A_{2s} \\ \vdots & \vdots & & \vdots \\ A_{s1} & A_{s2} & \ldots & A_{ss} \end{bmatrix} \qquad \sum_{i=1}^{s} n_i = n. \tag{4}$$

It follows easily that the operator norms $\| A \| = \sup_{x \neq 0} \| A\mathbf{x} \| / \| \mathbf{x} \|$ induced by the vector norms in (2) and (3) are

$$\| A \|_\infty = \max_{i \in S} \sum_{j \in S} \| A_{ij} \|_\infty$$

$$\| A \|_1 = \max_{j \in S} \sum_{i \in S} \| A_{ij} \|_1.$$

Lemma 2. Let the matrix A be partitioned as in (4), let A_k be the $n_k \times n$ matrix

$$A_k = [A_{k1}, A_{k2} \ldots A_{ks}] \qquad k = 1 \ldots s,$$

and assume that the (square) matrix A_{kk} is nonsingular. Then (a′) ⇒ (b′) where

(a′) $1 > \| A_{kk}^{-1} \|_\infty \sum_{j \neq k} \| A_{kj} \|_\infty$

(b′) $A_k\mathbf{x} = A_{k1}\mathbf{x}_1 + A_{k2}\mathbf{x}_2 + \ldots + A_{ks}\mathbf{x}_s = \mathbf{0}$, $\mathbf{x} \neq 0$ implies $\| \mathbf{x}_k \|_\infty < \| \mathbf{x} \|_\infty$ for all $\mathbf{x} \in R^n$ partitioned as in (1).

Proof: We have for each k, $1 \leq k \leq s$,

$$\begin{aligned} A_{kk}\mathbf{x}_k &= -\sum_{j \neq k} A_{kj}\mathbf{x}_j \\ \| A_{kk}\mathbf{x}_k \|_\infty &\leq \sum_{j \neq k} \| A_{kj}\mathbf{x}_j \|_\infty \\ &\leq \sum_{j \neq k} \| A_{kj} \|_\infty \| \mathbf{x}_j \|_\infty \\ &\leq \sum_{j \neq k} \| A_{kj} \|_\infty \| \mathbf{x} \|_\infty . \end{aligned} \tag{5}$$

Also,

$$\| x_k \|_\infty = \| A_{kk}^{-1} A_{kk}\mathbf{x}_k \|_\infty \leq \| A_{kk}^{-1} \|_\infty \| A_{kk}\mathbf{x}_k \|_\infty \tag{6}$$

$$\| A_{kk}\mathbf{x}_k \|_\infty \geq \frac{\| \mathbf{x}_k \|_\infty}{\| A_{kk}^{-1} \|_\infty} .$$

Combining (5) and (6), we obtain

$$\frac{\| \mathbf{x}_k \|_\infty}{\| A_{kk}^{-1} \|_\infty} < \sum_{j \neq k} \| A_{kj} \|_\infty \| \mathbf{x} \|_\infty$$

or

$$\| x_k \|_\infty \leq \| A_{kk}^{-1} \|_\infty \sum_{j \neq k} \| A_{kj} \|_\infty \| \mathbf{x} \|_\infty . \tag{7}$$

(7) combined with the assumption (a′) above gives

$$\| \mathbf{x}_k \|_\infty < \| \mathbf{x} \|_\infty .$$

whereas in the scalar case ($n_r = 1$, $s = n$) it is also true that (b′) $\Rightarrow$ (a′), see [7], in the matrix case in Lemma 2 in general (b′) $\not\Rightarrow$ (a′). However, the sufficiency of (a′) is enough for our purposes later.

Counterexample: Take $n = 4$, $s = 2$, $n_1 = 2$, $n_2 = 2$ and let

$$A_{11} = \begin{bmatrix} \frac{1}{10} & 0 \\ 0 & 40 \end{bmatrix} \qquad A_{12} = - \begin{bmatrix} \frac{1}{40} & \frac{1}{40} \\ 10 & 10 \end{bmatrix} .$$

Then

$$A_{11}\mathbf{x}_1 + A_{12}\mathbf{x}_2 = \mathbf{0}$$

becomes

$$\begin{bmatrix} \frac{1}{10} & 0 \\ 0 & 40 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix} - \begin{bmatrix} \frac{1}{40} & \frac{1}{40} \\ 10 & 10 \end{bmatrix} \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} = 0. \tag{8}$$

Now, (b′) holds because if $\mathbf{x}_1$ and $\mathbf{x}_2$ satisfy (8) and they are both not zero, then they satisfy (b′) since

$$\left.\begin{aligned} | x_{11} | &\leq \frac{1}{2} \| \mathbf{x}_2 \|_\infty \\ | x_{12} | &\leq \frac{1}{2} \| \mathbf{x}_2 \|_\infty . \end{aligned}\right\} \Rightarrow \| \mathbf{x}_1 \|_\infty < \| \mathbf{x}_2 \|_\infty = \| \mathbf{x} \|_\infty$$

However, (a′) does not hold because:

$$A_{11}^{-1} = \begin{bmatrix} 10 & 0 \\ 0 & \frac{1}{40} \end{bmatrix}, \qquad \| A_{11}^{-1} \|_\infty = 10$$

$$\| A_{12} \|_\infty = 20$$

and

$$\| A_{11}^{-1} \|_\infty \| A_{12} \|_\infty = (10)(20) = 200 > 1$$

Definition 2

An $n \times n$ matrix partitioned as in (4) is called blockwise (strictly) diagonally dominant if its kth row-block A_k is blockwise (strictly) diagonally dominant, i.e. if for each $k = 1, 2 \ldots s$

$$\| A_{kk}^{-1} \|_\infty \sum_{j \neq k} \| A_{kj} \|_\infty \leq 1. \qquad (<)$$

The following examples show that neither of the two notions for matrices, pointwise or blockwise strict diagonal dominance, implies the other.

Example (pointwise strict block diagonal dominance does not imply blockwise diagonal dominance). The matrix

$$A = \left[\begin{array}{cc|cc} 5 & 3 & 1 & 0 \\ 5 & 8 & 0 & 0 \\ \hline 0 & 1 & 3 & 1 \\ 7 & 0 & 0 & 8 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

is pointwise strictly diagonal dominant. However since

$$\| A_{22}^{-1} \|_\infty = \left\| \begin{bmatrix} \frac{1}{3} & -\frac{1}{24} \\ 0 & -\frac{1}{8} \end{bmatrix} \right\|_\infty = \frac{3}{8}$$

$$\| A_{21} \|_\infty = 7,$$

we have

$$\| A_{22}^{-1} \|_\infty \| A_{21} \| = \frac{21}{8} > 1$$

so that A is not blockwise diagonally dominant.

Example (blockwise strict diagonal dominance does not imply pointwise diagonal dominance). The matrix

$$A = \left[\begin{array}{cc|cc} 3 & 5 & 1 & 0 \\ 8 & 5 & 0 & 0 \\ \hline \frac{1}{2} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

is not pointwise diagonally dominant. However, since

$$\| A_{11}^{-1} \|_\infty \| A_{12} \|_\infty = \left\| \begin{bmatrix} -\frac{1}{5} & \frac{1}{5} \\ \frac{8}{25} & -\frac{3}{25} \end{bmatrix} \right\|_\infty \cdot 1 = \frac{11}{25} < 1$$

$$\| A_{22}^{-1} \|_\infty \| A_{21} \|_\infty = 1 \cdot \frac{1}{2} < 1$$

A is blockwise strictly diagonally dominant.

Although in the (vector) case the implication in Lemma 2 is in only one direction, whereas in the scalar case there is an equivalence, we nevertheless adopt the following definitions for block diagonal dominance for nonlinear mappings because they lead to sufficient conditions for convergence of iterations.

Let $\mathbf{x}, \mathbf{y} \in R^n$ be partitioned as in (1).

Definition 3

A function $\mathbf{f}_k$ partitioned into

$$\mathbf{f}_k^T = (f_{k1}^T, f_{k2}^T \dots f_{kn_k}^T) \qquad k = 1, 2 \dots s$$
$$\mathbf{f}_k: D \subset R^n \to R^{n_k}$$

is *blockwise* strictly diagonally dominant on D with respect to the kth block if for every $\mathbf{x} \neq \mathbf{y}$ in D

$$\mathbf{f}_k(\mathbf{x}) = \mathbf{f}_k(\mathbf{y}) \text{ implies that } \| \mathbf{x}_k - \mathbf{y}_k \|_\infty < \| \mathbf{x} - \mathbf{y} \|_\infty$$

Definition 4

A function F partitioned into

$$F^T = (\mathbf{f}_1^T, \mathbf{f}_2^T \dots \mathbf{f}_s^T)$$
$$F: D \subset R^n \to R^n \tag{9}$$

is *blockwise* strictly diagonally dominant on D if for each $k = 1 \dots s$ $\mathbf{f}_k$ is blockwise strictly diagonally dominant on D with respect to the kth block.

We now prove a result that gives a sufficient condition for a function to be blockwise diagonally dominant in terms of its derivative. As in [7] the notion of differentiability used is that of the Gateaux deirvative. We use the notation $\partial_i \mathbf{f}_k \equiv \partial \mathbf{f}_k / \partial \mathbf{x}_i$, which is an $n_k \times n_i$ matrix, and $f'_k \equiv (\partial_1 \mathbf{f}_k, \partial_2 \mathbf{f}_k \dots \partial_s \mathbf{f}_k)$.

THEOREM 2.1: Let $F: D \subset R^n \to R^n$ be G-differentiable on the convex set $D_0 \subset D$. Assume that for each $\mathbf{x}, \mathbf{y} \in D_0$ the matrix $A(\mathbf{x}, \mathbf{y}) = [A_{kj}(\mathbf{x}, \mathbf{y})]$, $1 \leq k, j \leq s$

$$A_{kj}(\mathbf{x}, \mathbf{y}) = \int_0^1 \partial_j \mathbf{f}_k[\mathbf{x} + t(\mathbf{y} - \mathbf{x})] \, dt \tag{10}$$

is a blockwise strictly diagonally dominant matrix. Then F is a blockwise strictly diagonally dominant function on D_0.

Proof: Let $k \in S$ be given and assume that $\mathbf{f}_k(\mathbf{x}) = \mathbf{f}_k(\mathbf{y})$ for some $\mathbf{x} \neq \mathbf{y}$ in D_0. Applying the integral form of the Mean Value Theorem we have

$$\int_0^1 \mathbf{f}_k' [\mathbf{x} + t(\mathbf{y} - \mathbf{x})] (\mathbf{y} - \mathbf{x})\, dt = 0$$

or

$$\sum_{j=1}^{s} \int_0^1 \partial_j \mathbf{f}_k[\mathbf{x} + t(\mathbf{y} - \mathbf{x})]\, dt\, (\mathbf{y}_j - \mathbf{x}_j) = \mathbf{0}$$

that is

$$\sum_{j=1}^{s} A_{kj}(\mathbf{y}_j - \mathbf{x}_j) = \mathbf{0}$$

Using Lemma 2 this implies that

$$\| \mathbf{y}_j - \mathbf{x}_j \|_\infty < \| \mathbf{y} - \mathbf{x} \|_\infty \quad \text{for} \quad j = 1, 2 \ldots s$$

which is the criterion for F to be a blockwise strict diagonally dominant function on D_0.

The integral in (10) may be difficult to evaluate in practice. If F is a quadratic mapping, however, (10) can be replaced by an expression which is simpler to evaluate.

THEOREM 2.2: Let F: $D \subset R^n \to R^n$ be a quadratic mapping on the convex set $D_0 \subset D$: that is, let f_{ki} denote the ith component of $\mathbf{f}_k$ then for $1 \leq k \leq s$, $1 \leq i \leq n_k$, $1 \leq r, t \leq n$

$$f_{ki}(\mathbf{x}) = \sum_{r,t} a_{rt}^{ki}\, x_r x_t + \sum_r b_r^{ki}\, x_r + c^{ki}. \tag{11}$$

Suppose that the matrix $F'(\mathbf{x})$ is a blockwise strictly diagonally dominant matrix for each $\mathbf{x} \in D_0$. Then F is a blockwise strictly diagonally dominant function on D_0.

Proof: From (11) we have for $1 \leq m \leq n$

$$(\partial/\partial x_m) f_{ki}(\mathbf{x}) = \sum_r (a_{rm}^{ki} + a_{mr}^{ki}) x_r + b_m^{ki}.$$

The (i, m)th element of the matrix A_{kj} in (10) is of the form

$$\begin{aligned}
&\int_0^1 \partial_m f_{ki}[\mathbf{x} + t(\mathbf{x} - \mathbf{y})]\, dt \\
&= \int_0^1 \left\{ \sum_r (a_{rm}^{ki} + a_{mr}^{ki})[x_r + t(y_r - x_r)] + b_m^{ki} \right\} dt \\
&= \sum_r (a_{rm}^{ki} + a_{mr}^{ki}) \left[\frac{1}{2}(x_r + y_r) \right] + b_m^{ki} \\
&= (\partial/\partial x_m) f_{ii} \left[\frac{1}{2}(\mathbf{x} + \mathbf{y}) \right].
\end{aligned}$$

Therefore in (10) we now have

$$A_{kj} = \partial_j \mathbf{f}_k \left[\frac{1}{2} (\mathbf{x} + \mathbf{y}) \right] .$$

Therefore, the criterion for $A_{kj}(\mathbf{x}, \mathbf{y})$ to be a blockwise strictly diagonally dominant matrix for *all* $\mathbf{x}, \mathbf{y} \in D_0$ becomes blockwise strict diagonally dominance for the matrix $F'(\mathbf{x})$.

In many applications, including the load flow problem for electric power systems, the matrices involved are block diagonally dominant but not strictly block diagonally dominant. In such cases we introduce a generalization of the notion of block irreducibility introduced in [3]:

Definition: An $n \times n$ matrix A partitioned as in (4) is block irreducible if the $s \times s$ matrix $B = (b_{ij})$ where $b_{ij} = \| A_{ij} \|$ is irreducible, i.e. if the graph of B is strongly connected.

Following the development in [7] we first introduce a generalization of a block strictly diagonally dominant matrix by stating a lemma. The proof is similar to that of Lemma 2.

Lemma 3: Let the matrix A be partitioned as in (4). Let A_k be the $n_k \times n$ matrix

$$A_k = [A_{k1}, A_{k2} \ldots A_{ks}] \qquad k = 1 \ldots s$$

and assume that the square matrix A_{kk} is nonsingular. Then (a′) $\Rightarrow$ (b′) where

(a′) $1 \geq \| A_{kk}^{-1} \|_\infty \sum_{j \neq k} \| A_{kj} \|_\infty$

(b′) $A\mathbf{x} = A_{k1}\mathbf{x}_1 + A_{k2}\mathbf{x}_2 + \cdots + A_{kS}\mathbf{x}_S = \mathbf{0}, \mathbf{x} \neq 0$

implies that either $\| \mathbf{x}_k \|_\infty < \| \mathbf{x} \|_\infty$ or $\| \mathbf{x}_k \|_\infty = \| \mathbf{x} \|_\infty = \| x_j \|_\infty$, whenever $\| A_{kj} \|_\infty \neq 0$ for any $\mathbf{x} \in R^n$, partitioned as in (1).

In order to specify the counterpart of the condition $\| A_{kj} \|_\infty \neq 0$ for nonlinear mappings we extend to the block case the definition of diagonal dominance with respect to a family of networks which was introduced in [7]. A *network* $\Omega = (N, \Lambda)$ consists of a set of n *nodes* $N = \{1, \ldots, n\}$, and a set $\Lambda \subset N \times N$ of directed *links* which contain no loops; that is, $(i, i) \in \Lambda$ if $i \in N$. *A node i is connected to a node j* if there is a directed path in Λ from i to j; that is, a sequence of links of the form $(i, i_1), (i_1, i_2) \ldots (i_r, j)$.

Definition 5

Let $\mathbf{x}, \mathbf{y} \in R^n$ be partitioned as in (1) and let F: $D \subset R^n \rightarrow R^n$ be partitioned as in (9). The mapping F is blockwise diagonally dominant on D with respect to the family of networks $\{\Omega_{x,y}$: $\mathbf{x}, \mathbf{y} \in D, \mathbf{x} \neq \mathbf{y}\}$ if for every $\mathbf{x}, \mathbf{y} \in D, \mathbf{x} \neq \mathbf{y}$ the network $\Omega_{x,y} = (S, \Lambda_{x,y})$ is such that $\mathbf{f}_k(\mathbf{y}) = \mathbf{f}_k(\mathbf{x})$ and $k \in S$ implies that either $\| \mathbf{x}_k - \mathbf{y}_k \|_\infty < \| \mathbf{x} - \mathbf{y} \|_\infty$ or $\| \mathbf{x}_k - y_k \|_\infty = \| \mathbf{x} - \mathbf{y} \|_\infty = \| \mathbf{x}_j - \mathbf{y}_j \|_\infty$ whenever $(k, j) \in \Lambda_{x,y}$.

If follows from Lemma 3 that for F a linear mapping and A the associated matrix, if A is a blockwise diagonally dominant matrix then F is a blockwise diagonally dominant mapping on all of R^n with respect to the network $\Omega_A = (S, \Lambda_A)$. Here Ω_A is the same for all x, y, and Λ_A is defined as

$$\Lambda_A = \{(i, j) \in S \times S, \| A_{ij} \|_\infty \neq 0\}$$

The converse is, however, not true, as illustrated by the following counterexample:

Counterexample: let $n = 4$, $s = 2$, $n_1 = 2$, $n_2 = 2$, A_{11}, A_{12} be defined as in (8), and

$$A_{21} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad A_{22} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

We have already shown that $\mathbf{f}_1(\mathbf{y}) = \mathbf{f}_1(\mathbf{x})$ for some $\mathbf{y} \neq \mathbf{x}$ implies that $\| \mathbf{x}_1 - \mathbf{y}_1 \|_\infty < \| \mathbf{x} - \mathbf{y} \|_\infty$. If $f_2(\mathbf{y}) = f_2(\mathbf{x})$ for $\mathbf{y} \neq \mathbf{x}$ then $\| \mathbf{x}_2 - \mathbf{y}_2 \|_\infty = \frac{1}{2} \| \mathbf{x}_1 - \mathbf{y}_1 \|_\infty < \| \mathbf{x} - \mathbf{y} \|_\infty$ so that F is a blockwise diagonally dominant mapping for all $\mathbf{x}, \mathbf{y} \in R^n$, $\mathbf{x} \neq \mathbf{y}$ with respect to Ω_A. In this case $\Lambda_A = \{(1, 2), (2, 1)\}$. As we have shown, however, by considering the first row block A is not a blockwise diagonally dominant matrix.

In the case when all the blocks are one by one (that is, $n_i = 1$ for all i) and $\Omega_{x,y}$ does not depend upon y, then the definition reduces to that in [7] for a (pointwise) diagonally dominant mapping on D with respect to a family of networks $\{\Omega_x : \mathbf{x} \in D\}$. In the pointwise case (b′) also implies (a′). Also, there is then an equivalence for linear maps between a diagonally dominant matrix A and a diagonally dominant mapping with respect to Ω_A.

Next we formulate the concept of a block irreducible matrix as a special case of a broader definition for general maps.

Definition 6

Let $\mathbf{x}, \mathbf{y} \in R^n$ be partitioned as in (1) and let $F: D \subset R^n \to R^n$ be partitioned as in (9). The mapping F is blockwise weakly Ω-diagonally dominant on D if for every $\mathbf{x}, \mathbf{y} \in D$, $\mathbf{x} \neq \mathbf{y}$, there is a network $\Omega_{x,y} = (S, \Lambda_{x,y})$ such that:

(a) F is blockwise diagonally dominant with respect to the family of networks $\{\Omega_{x,y}\}$, and
(b) for every $\mathbf{x}, \mathbf{y} \in D$, $\mathbf{x} \neq \mathbf{y}$ there is a nonempty subset $J_{x,y}$ of S such that for each $i \in J_{x,y}$, $\mathbf{f}_i(\mathbf{y}) = \mathbf{f}_i(\mathbf{x})$ implies that $\| \mathbf{y}_i - \mathbf{x}_i \|_\infty < \| y - \mathbf{x} \|_\infty$ and for each $i \notin J_{x,y}$ there is a path in $\Lambda_{x,y}$ from i to some $j = j(i) \in J_{x,y}$. Sometimes the networks Ω_x and the sets J_x are independent of $\mathbf{x} \in D$.

Definition 7

Let $\mathbf{x} \in R^n$ be partitioned as in (1) and let $F: D \subset R^n \to R^n$ be partitioned as in (9). The mapping F is blockwise Ω-diagonally dominant on D if there is a network $\Omega = (S, \Lambda)$ such that

(a) F is blockwise diagonally dominant with respect to the network $\Omega = (S, \Lambda)$, and
(b) there is a nonempty subset J of S such that for each $i \in J$, $\mathbf{f}_i$ is a strictly diagonally dominant mapping with respect to $\mathbf{x}_i$ and for each $i \notin J$ there is a path in Λ from i to some $j = j(i) \in J$.

In particular, if F is a linear mapping on all of R^n and A is its associated matrix, then A (weakly) blockwise Ω-diagonally dominant on R^n means that F is (weakly) blockwise Ω-diagonally dominant on R^n with $\Omega_{x,y} = \Omega_A$ for all $\mathbf{x}, \mathbf{y} \in R^n$, $\mathbf{x} \neq \mathbf{y}$. In this case the following theorem is an immediate consequence of the definitions and Lemma 3.

THEOREM 2.3: Let A be an $(n \times n)$ matrix partitioned as in (4). Suppose that A is a blockwise diagonally dominant matrix. Then A is blockwise weakly Ω-diagonally dominant if and only if A is blockwise Ω-diagonally dominant.

The condition on J in (b) is a generalization of block irreducibility for a matrix, and the definition of a blockwise Ω-diagonally dominant mapping is a generalization of a blockwise irreducibly diagonally dominant matrix (see [3]).

Now we prove the analogs of Theorems 2.1 and 2.2 for Ω-blockwise diagonally dominant matrices.

THEOREM 2.4: Let $F: D \subset R^n \rightarrow R^n$ be continuously differentiable on the convex set $D_0 \subset D$. Assume that for each $\mathbf{x}, \mathbf{y} \in D_0$ the matrix $A(\mathbf{x}, \mathbf{y})$ defined in (10) is an Ω-blockwise diagonally dominant matrix. Then F is a blockwise weakly Ω-diagonally dominant function on D_0.

Proof: Let $\mathbf{x}, \mathbf{y} \in D_0$, $\mathbf{x} \neq \mathbf{y}$ be given, set $A = A(\mathbf{x}, \mathbf{y})$, $\Lambda_{x,y} = \{(i, j);\ i, j \in S,\ i \neq j,\ \| A_{ij} \|_\infty \neq 0\}$, $\Omega_{x,y} = (S, \Lambda_{x,y})$, $J_{x,y} = \{i\!: 1 > \| A_{ii}^{-1} \|_\infty \sum_{j \neq i} \| A_{ij} \|_\infty\}$. Then for each $i \notin J_{x,y}$, there is a path in $\Lambda_{x,y}$ from i to some $j = j(i) \in J_{x,y}$. We now show that F is blockwise diagonally dominant on D_0 with respect to the family of networks $\{\Omega_{x,y}\}$. Suppose for $k \in S$ $\mathbf{f}_k(\mathbf{y}) = \mathbf{f}_k(\mathbf{x})$, and $\| \mathbf{y}_k - \mathbf{x}_k \|_\infty = \| \mathbf{y} - \mathbf{x} \|_\infty$. Applying the integral form of the Mean Value theorem as in the proof of Theorem 2.1, we again have

$$\sum_{j=1}^{s} A_{kj}(\mathbf{y}_j - \mathbf{x}_j) = \mathbf{0}.$$

Since by assumption A is an Ω blockwise diagonally dominant matrix, it follows from Lemma 3 that $\| \mathbf{y}_j - \mathbf{x}_j \|_\infty = \| \mathbf{y} - \mathbf{x} \|_\infty$ whenever $(k, j) \in \Omega_{x,y}$. It also follows from Lemma 2 that if $k \in J_{x,y}$ then $\| \mathbf{y}_k - \mathbf{x}_k \|_\infty < \| \mathbf{y} - \mathbf{x} \|_\infty$. This concludes the proof.

THEOREM 2.5: Let $F: D \subset R^n \rightarrow R^n$ be a quadratic mapping on the convex set $D_0 \subset D$ defined as in (11). Suppose that the matrix $F'(\mathbf{x})$ is an Ω-blockwise diagonally dominant matrix for each $\mathbf{x} \in D_0$. Then $F(\mathbf{x})$ is a blockwise Ω weakly diagonally dominant function on D_0.

Proof: The proof closely parallels that of Theorem 2.2.

3. UNDERRELAXED BLOCK JACOBI AND BLOCK GAUSS-SEIDEL ITERATIONS

Definition 8 (Underrelaxed Blockwise Jacobi Iterative method).

Let $F: D \subset R^n \rightarrow R^n$ be partitioned as in (9), and let $\mathbf{x} \in R^n$ be partitioned as in (1).

Then the underrelaxed Block Jacobi iterative method is as follows: given an initial guess $\mathbf{x}^{(0)} \in R^n$, for each $k \in S$, solve for $\mathbf{x}_k$ from $\mathbf{f}_k$

$$\mathbf{f}_k(\mathbf{x}_1^p, \mathbf{x}_2^p \ldots \mathbf{x}_{k-1}^p, \mathbf{x}_k, \mathbf{x}_{k+1}^p \ldots \mathbf{x}_s^p) = 0$$

and take

$$\mathbf{x}_k^{p+1} = (1 - \omega)\mathbf{x}_k^p + \omega \mathbf{x}_k \tag{12}$$

for a given under relaxation parameter

$$0 < \omega \leq 1, \quad \text{and for} \quad p = 0, 1, 2 \ldots.$$

Definition 9 (Underrelaxed Blockwise Gauss-Seidel Iterative Method).

Let F: $D \subset R^n \to R^n$ be partitioned as in (9) and let $\mathbf{x} \in R^n$ be partitioned as in (1). Then the underrelaxed Block Gauss-Seidel iterative method is as follows: given initial guesses $\mathbf{x}_k^{(0)}$, $k = 1, 2 \ldots s$ solve successively for $\mathbf{x}_k$ $k = 1, 2 \ldots s$ from

$$\mathbf{f}_k(\mathbf{x}_1^{p+1}, \mathbf{x}_2^{p+1} \ldots \mathbf{x}_{k-1}^{p+1}, \mathbf{x}_k, \mathbf{x}_{k+1}^p \ldots \mathbf{x}_s^p) = \mathbf{0}$$

and take

$$\mathbf{x}_k^{p+1} = (1 - \omega)\mathbf{x}_k^p + \omega \mathbf{x}_k \tag{13}$$

for a given under relaxation parameter $0 < \omega \leq 1$, and for $p = 0, 1, 2 \ldots$.

We now derive sufficient conditions for the underrelaxed Block Jacobi and Block Gauss-Seidel iterations to converge. Our development parallels that of Moré [7].

Definition 10

A rectangle Q in R^n is the Cartesian product of n intervals I_i, each of which may be either open, closed, or semi-open. (The form $(\alpha_i, +\infty)$ or $[\alpha_i, +\infty)$; $\alpha_i = -\infty$ is permitted in the first form; otherwise, α_i is real.)

$$Q = \prod_{i=1}^{n} I_k$$

THEOREM 3.1: Let Q be a convex set in the case of Block Gauss-Seidel iterations or let Q be a rectangle in the case of Block Jacobi iterations. Let F: $Q \subset R^n \to R^n$ be partitioned as in (9), suppose that F is a blockwise weakly diagonally dominant function on Q, and suppose that for each $\mathbf{x} \in Q$ and $k \in S$, then n_k equations

$$\mathbf{f}_k(\mathbf{x}_1, \mathbf{x}_2 \ldots \mathbf{x}_{k-1}, \mathbf{t}, \mathbf{x}_{k+1} \ldots \mathbf{x}) = 0$$

have a unique solution $\mathbf{t}_k = [t_{k1}^*, t_{k2}^*, \ldots, t_{kn_k}^*] \in Q$. Then the Block Gauss-Seidel and Block Jacobi sequences defined as in (12) and (13) are well defined for *any* $\mathbf{x}^0 \in Q$ and for either method there is an iteration function H: $Q \subset R^n \to R^n$ such that:

a) The method is equivalent to $\mathbf{x}^{k+1} = H\mathbf{x}^k$, $k = 0, 1 \ldots$,
b) $H(Q) \subset Q$,
c) $\| H\mathbf{x} - H\mathbf{y} \|_\infty \leq \| \mathbf{x} - \mathbf{y} \|_\infty$ for every $\mathbf{x}, \mathbf{y} \in Q$.

Moreover, if $F\mathbf{x} = 0$ has a unique solution $\mathbf{x}$ in Q, then

d) $\| H^{n-l-1} \mathbf{x} - \mathbf{x}^* \|_\infty < \| \mathbf{x} - \mathbf{x}^* \|_\infty$ for every $\mathbf{x} \neq \mathbf{x}^*$ where l denotes the number of elements in $J_{x,x}^*$.

Proof: The proof is exactly the same as that of Theorem 5.3 in [7] with $\| \mathbf{h}_i(\mathbf{x}) - \mathbf{h}_i(y) \|_\infty$ replacing $| h_i(\mathbf{x}) - h_i(y) |$ and $\Lambda_{x,y}$ replacing Λ_x where appropriate. We do not repeat the proof here. The next corollary can be proved in the same way.

Corollary 3.2: Let F: $Q \subset R^n \to R^n$ satisfy the hypotheses of Theorem 3.1 on the closed rectangle Q. Then $F\mathbf{x} = 0$ has a unique solution $\mathbf{x}^* \in Q$ if and only if for some $x^0 \in Q$ the Jacobi or Gauss-Seidel iterates in (12) and (13) with $\omega \in (0, 1]$ are bounded. In particular this occurs if Q is bounded; in any case, the iterates will converge to x^* for any $x^0 \in Q$.

Our principal concern is the application of Corollary 3.2 to the solution of the load flow problem for electric power systems. We now briefly describe a new formulation of the load flow problem which leads to a system of nonlinear equations $F\mathbf{x} = 0$ where each $f_k(\mathbf{x})$ is a quadratic function.

4. THE $\hat{S}$–$\hat{E}$ MODEL OF THE LOAD FLOW PROBLEM FOR ELECTRIC POWER SYSTEMS

Every electric power system consists of a set of

(i) buses,
(ii) generators,
(iii) loads,
(iv) transmission lines.

In [6] a new mathematical model is formulated for an electric power system. In this model the unknowns are complex valued voltages (denoted by $\hat{E}$). Complex valued power flows (denoted by $\hat{S}$) are described in terms of the voltages $\hat{E}$ and a generalized Kirchhoff power flow law is applied.

An $\hat{S}$–$\hat{E}$ graph is used to represent the electric power system. In this graph the *buses* are represented by *nodes* and $\hat{E}_i$ is the voltage at bus i. Each *generator* is represented by a *single branch* and associated with this branch is the (complex) value of an ideal power flow source called an input or an injection. Each *load* is also represented by a *single branch* and associated with this branch is the (complex) value of another ideal power flow, again called an input or an injection. Each *transmission line*, however, is represented by *two branches*: associated with one branch is the (complex) value of the *transmitted* power flow from one end of the line to the other, i.e. from one bus to another (the transmitted power flow is a known (quadratic) function of the voltages at both ends of the line), and associated with the other branch is the (complex) value of power *loss* flow from one end of the line to the other (the power loss flow is another known (quadratic) function of the voltages at both ends of the line). The concept of representing transmission lines by *two* branches of the graph is new and is an essential contribution to the formulation in [6]. Each *generator* i has a fixed injected active power P_i into node i and a fixed known magnitude V_i of the complex voltage at node i (often called a PV bus). Each *load* i has fixed known injected active power P_i and reactive power Q_i into node i (often called PQ bus). Finally there is always at least one bus called the "*slack bus*", which has a fixed known (complex) voltage. In the $\hat{S}$–$\hat{E}$ graph a slack bus is always connected to a single branch and associated with this branch is an ideal voltage source with the same value as the voltage of the slack bus.

An example of a small electric power system is shown in Fig. 1 in the form of a one-line diagram [2].

The simplest way of describing this system is that it consists of: one generator, the branch connecting node 1 to the ground (node 1, in this case, is also taken to be the slack bus); two loads between each of the nodes $i = 2, 3$, and the ground; and three transmission lines, from 1 to 2, from 2 to 3, and from 1 to 3. Each node $i = 2, 3$, has a known load associated with it. This load (input) is denoted by $S_i = P_i + \sqrt{-1}\, Q_i$. The voltages $\hat{E}_i$ $i = 2, 3$ are to be determined. The $\hat{E}_i$ represent the steady-state voltage differences between the voltage at the ith node and the ground. A physical constraint is that for each i, $Re\ \hat{E}_i > 0$.

For our purposes we represent each complex number $\hat{E}_i$ using rectangular coordinates:

$$\hat{E}_i = E_{ix} + \sqrt{-1}\, E_{iy} \qquad E_{ix} > 0 \qquad i = 1, 2, 3 \ldots .$$

The voltage of the ground is considered to have magnitude zero.

Each transmission line connecting node i and node k is characterized by its complex impedance $\hat{Z}_{ik} = R_{ik} + \sqrt{-1}X_{ik}$ where R_{ik} is the resistance and X_{ik} the reactance, or equivalently by its complex admittance $\hat{Y}_{ik} = 1/\hat{Z}_{ik} = G_{ik} - \sqrt{-1}B_{ik}$ where G_{ik} is the conductance and B_{ik} is the susceptance. R_{ik} and G_{ik} correspond to real power losses on the line. In particular for lossless lines $\hat{Z}_{ik} = \sqrt{-1}\,X_{ik}$, $\hat{Y}_{ik} = -\sqrt{-1}\,B_{ik}$. B_{ik} is always positive.

The $\hat{S}$–$\hat{E}$ graph corresponding to Fig. 1 is shown in Fig. 2. Completely transmitted power from node i to node k, $\hat{S}^T_{ik}$, has the form

$$\hat{S}^T_{ik} = \frac{\hat{E}_k(\hat{E}_i - \hat{E}_k)^*}{\hat{Z}^*_{ik}} = \hat{E}_k(\hat{E}_i - \hat{E}_k)^*\,\hat{Y}^*_{ik}.$$

* denotes complex conjugation. Total power loss on this line because of the flow from node i to node k, $\hat{S}^L_{ik}$, has the form

$$\hat{S}^L_{ik} = \frac{|\,\hat{E}_i - \hat{E}_k\,|^2}{\hat{Z}^*_{ik}} = |\,\hat{E}_i - \hat{E}_k\,|^2\,\hat{Y}^*_{ik}.$$

A basic property of an $\hat{S}$–$\hat{E}$ graph is that power flows (instead of currents) satisfy Kirchhoff's flow law: the sum of power flows at each node is zero. For more details see [6].

Assume that there are $(n + 1)$ nodes in the $\hat{S}$–$\hat{E}$ graph. The slack node is taken to be the $(n + 1)$st node. Let all the loads of a given system belong to a set $\{PQ\}$ and all the generators to a set $\{PV\}$. Then, given the values of the inputs S_i at all PQ buses, the active power injections P_i and voltage magnitudes V_i of all PV buses of an electric power

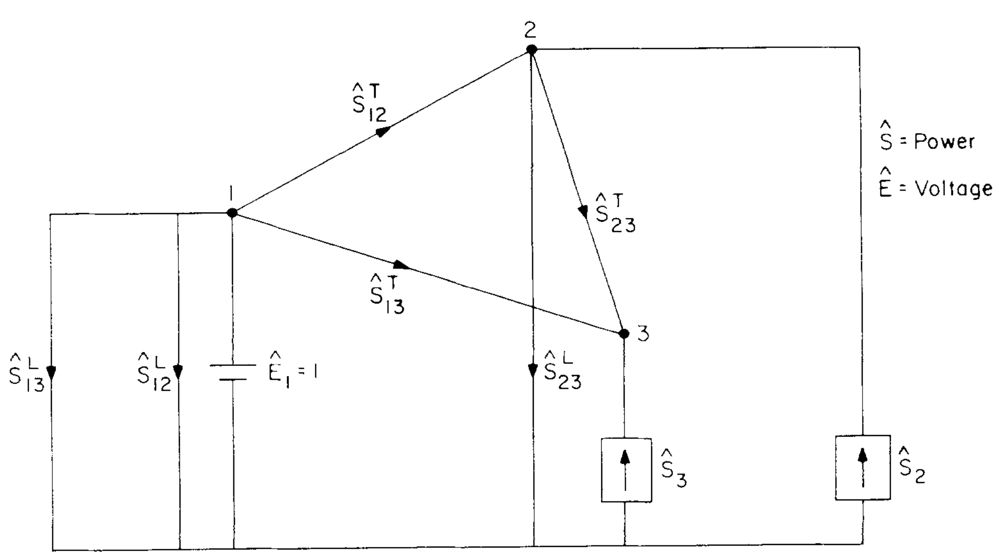

Fig. 1. One-line diagram for a three-bus system

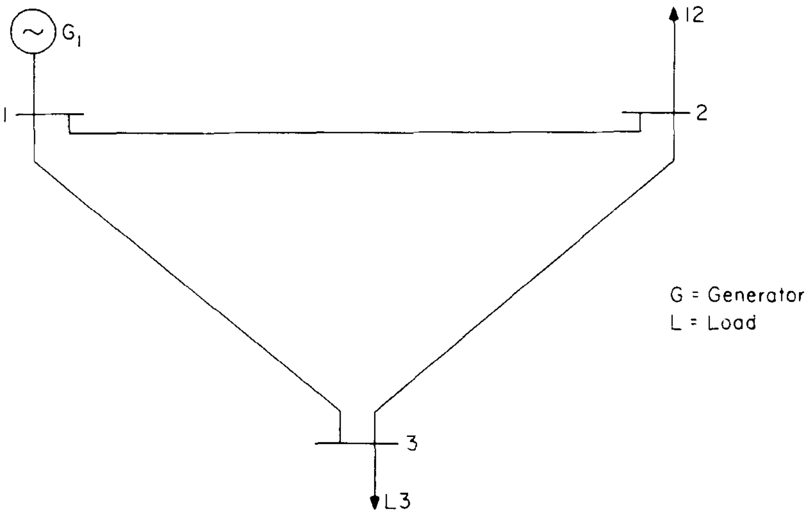

Fig. 2. S–E graph for a three-bus system.

system, the load flow problem is to find the real and imaginary parts of the voltages at the buses. The voltage at the slack bus is assumed to be known.

There is currently no closed form solution for this problem. All the algorithms are of iterative type; a detailed overview is given in [10].

The algorithm whose convergence properties we analyze is given in detail in [6]. The working equations of this algorithm depend on the choice of a minimal spanning tree for a given $\hat{S}$–$\hat{E}$ graph. In general there are many choices for the minimal spanning tree.

We analyze the algorithm when the minimal spanning tree has a single root which is taken to be the ground. For this choice the implicit equations describing voltages in terms of injections have the same form as the conventional load flow equations [10].

For node j, let

- $\{k_j\}$ be the set of all nodes connected to node j which have transmitted power branches going *towards* j
- $\{l_j\}$ be the set of all nodes connected to node j which have transmitted power going *away* from j

It is always true that $\{k_j\} \cap \{l_j\} = \phi$.

Then the application of Kirchhoff's flow law leads to the following equation for $1 \leq j \leq n$:

$$\text{flow into node } j - \text{flow out of node } j = 0$$

$$\sum_{k\in\{k_j\}} \hat{S}^T_{kj} - \sum_{l\in\{l_j\}} \hat{S}^L_{jl} - \sum_{l\in\{l_j\}} \hat{S}^T_{jl} + \hat{S}_j = 0$$

$$\sum_{k\in\{k_j\}} \hat{E}_j(\hat{E}_k - \hat{E}_j)^* Y^*_{kj} - \sum_{l\in\{l_j\}} [|\hat{E}_j - \hat{E}_l|^2 + \hat{E}_l(\hat{E}_j - \hat{E}_l)^*] Y^*_{lj} + \hat{S}_j = 0.$$

$$j \in \{PQ\} \cup \{PV\}$$

This reduces to

$$\sum_{k\in\{k_j\}} \hat{E}_j(\hat{E}_k - \hat{E}_j)^* \hat{Y}^*_{kj} - \sum_{l\in\{l_j\}} \hat{E}_j(\hat{E}_j - \hat{E}_l)^* \hat{Y}^*_{jl} + \hat{S}_j = 0. \tag{14}$$

For lossless transmission lines, we have

$$Y_{kj} = -\sqrt{-1}\, B_{kj}, \quad Y_{jl} = -\sqrt{-1}\, B_{jl}, \quad \text{with} \quad B_{kj} > 0, \quad B_{jl} > 0$$

and (14) becomes for $1 \leq j \leq n$

$$-\sqrt{-1}\,\hat{f}_j = \sqrt{-1} \sum_{k \in \{k_j\}} \hat{E}_j(\hat{E}_k - \hat{E}_j)^* B_{kj} - \sqrt{-1} \sum_{l \in \{l_j\}} \hat{E}_j(\hat{E}_j - \hat{E}_l)^* B_{jl} + \hat{S}_j = 0 \tag{15}$$

$$j \in \{PQ\} \cup \{PV\}.$$

For lossless systems in particular, (15) is a quadratic system of n complex valued equations $\hat{f}_j = 0$ in the $2n$ real valued unknowns. Writing $\hat{f}_j = 0 = f_j^{(1)} + \sqrt{-1}\, f_j^{(2)}$, $\hat{S}_j = P_j + \sqrt{-1}Q_j$, we obtain $2L$ (L is the number of load buses) real valued equations for the load nodes

$$f_j^{(1)} = \sum_{k \in \{k_j\}} (E_{jx}E_{kk} + E_{jy}E_{ky})B_{kj} - (E_{jx}^2 + E_{jy}^2)B_{kj} + \sum_{l \in \{l_j\}} (E_{jx}E_{lx} + E_{jy}E_{ly})B_{jl} - (E_{jx}^2 + E_{jy}^2)B_{jl} + Q_j = 0 \tag{16a}$$

$$f_j^{(2)} = \sum_{k \in \{k_j\}} (E_{jy}E_{kx} - E_{jx}E_{ky})B_{kj} + \sum_{l \in \{l_j\}} (E_{jy}E_{lx} - E_{jx}E_{ly})B_{jl} - P_j = 0$$

$$j \in \{PQ\}. \tag{16b}$$

For the generator nodes, the voltage magnitudes instead of the injected reactive power are known so we have $2(n - L)$ real valued equations:

$$f_j^{(1)} = E_{jx}^2 + E_{jy}^2 - V_j^2 = 0 \tag{16c}$$

$$f_j^{(2)} = \sum_{k \in \{k_j\}} (E_{jy}E_{kx} - E_{jx}E_{ky})B_{kj} + \sum_{l \in \{l_j\}} (E_{jy}E_{lx} - E_{jx}E_{ly})B_{jl} - P_j = 0$$

$$j \in \{PQ\}. \tag{16d}$$

Let $\mathbf{f}_j = (f_j^{(1)}, f_j^{(2)})^T$, $\mathbf{x}_j = (E_{jx}, E_{jy})^T$. Then (16) is partitioned in a natural way with each n_i in (1) equal to 2, n in (1) replaced by $(2n)$ and s in (1) equal to n. Similarly each k_i in (9) is equal to 2 and $F = (\mathbf{f}_1^T, \mathbf{f}_2^T \ldots \mathbf{f}_n^T)^T$.

5. APPLICATION OF Ω BLOCK DIAGONAL DOMINANCE TO THE LOAD FLOW PROBLEM

We restrict the discussion in this section to a class of power systems which has the lossless property and which has no more than one generator (with *any* number of loads). We wish to use the block Jacobi and block Gauss-Seidel iterations (BGS) defined in (12) and (13) to solve (16). By Corollary 3.2 the iterations converge if there is a closed rectangle $Q \subset R^{2n}$ such that the following conditions are satisfied:

1) F is a blockwise Ω weakly diagonally dominant function on Q,
2) either Q is bounded or for some x^0 in Q the iterates are all bounded,

3) for each $k = 1, 2 \ldots n$,

$$\mathbf{f}_k(\mathbf{x}_1, \mathbf{x}_2 \ldots \mathbf{x}_{k-1}, \mathbf{t}, \mathbf{x}_{k+1} \ldots \mathbf{x}_n) = 0$$

has a unique solution $\mathbf{t}_k^* = [t_{k1}, t_{k2}]^T \in Q$.

We study each of these sufficient conditions when applied to the solution of $F\mathbf{x} = 0$ in (16) separately.

Conditions (1) *and* (2):

By Theorem 2.5 since F is a quadratic function, a sufficient condition for F to be a blockwise Ω weakly diagonally dominant function is that $F'(\mathbf{x})$ is an Ω blockwise diagonally dominant matrix for each $\mathbf{x} \in Q$. Let

$$A_{ji}(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f_j^{(1)}}{\partial E_{ix}} & \dfrac{\partial f_j^{(1)}}{\partial E_{iy}} \\ \dfrac{\partial f_j^{(2)}}{\partial E_{ix}} & \dfrac{\partial f_j^{(2)}}{\partial E_{iy}} \end{bmatrix}$$

$F'(\mathbf{x})$ is a blockwise diagonally dominant matrix if for $j = 1 \ldots n$ we have $\| A_{jj}^{-1} \|_\infty \sum_{i \neq j} \| A_{ji} \| \leq 1$.

It is generally assumed in electric power systems (and has been verified empirically) that one set of voltages called the flat start is close to the true (physical) solution to (16) (in the case under discussion only (16a), (16b) apply) which we denote by $\hat{E}'$. The flat start at each node is the voltage which has zero phase angle and magnitude 1; that is, denoting the flat start at node i by E_i^{FS} we have

$$E_{ix}^{FS} = 1, \qquad E_{iy}^{FS} = 0 \qquad i = 1 \ldots n.$$

We now show that the $2n \times 2n$ matrix $F'(\hat{E}^{FS})$ is blockwise diagonally dominant. It is easily calculated that at $\mathbf{x} = \hat{E}^{FS}$,

$$\| A_{jj}^{-1} \|_\infty = 1 \Big/ \left(\sum_{k \in \{k_j\}} B_{kj} + \sum_{l \in \{l_j\}} B_{jl} \right) \tag{17}$$

$$\| A_{jk} \|_\infty = B_{kj} \quad \text{if} \quad k \in \{k_j\}, \qquad k \neq n + 1 \tag{18}$$

$$\| A_{jl} \|_\infty = B_{jl} \quad \text{if} \quad l \in \{l_j\}, \qquad l \neq n + 1. \tag{19}$$

If the jth node is not connected to the $(n + 1)$st node (slack node) then $(n + 1) \notin \{k_j\}$ and $(n + 1) \notin \{l_j\}$. Therefore,

$$\| A_{jj}^{-1} \|_\infty \sum_{i \neq j} \| A_{ji} \|_\infty = \left(\sum_{k \in \{k_j\}} B_{kj} + \sum_{l \in \{l_j\}} B_{jl} \right) \Big/ \left(\sum_{k \in \{k_j\}} B_{kj} + \sum_{l \in \{l_j\}} B_{jl} \right) = 1. \tag{20}$$

There is always at least one node connected to the slack node. If j is such a node, then $(n + 1) \in \{k_j\} \cup \{l_j\}$ so that either $B_{n+1,j}$ or $B_{j,n+1}$ (or both) are missing from the sum in the numerator in (20). $B_{n-1,j}$ and $B_{j,n-1}$ are, however, both present in the sums in (17). Therefore, for any node connected to a slack node, strict inequality holds in (20):

$$\| A_{jj}^{-1} \|_\infty \sum_{i \neq j} \| A_{ji} \|_\infty < 1$$

It is not difficult to show that for any $\mathbf{x}$, the matrix $A(\mathbf{x}) = [A_{jk}(\mathbf{x})]$ is block irreducible. Therefore, we have shown that $F'(\hat{E}^{FS})$ is an Ω blockwise diagonally dominant matrix.

Since the physical solution $\hat{E}'$ to (16) is known to be close to $\hat{E}^{FS}$, *it is assumed* that there is a closed bounded rectangle Q containing $\hat{E}'$ such that $F'(\hat{E})$ is an Ω blockwise diagonally dominant matrix for all $\hat{E} \in Q$. Under this assumption, using Theorem 2.5, conditions (1) and (2) are satisfied.

Condition (3):

In order to satisfy condition (3), $\mathbf{E}_j = (E_{jx}, E_{jy})^t$ must be solved for, in (16) for $j = 1, 2 \ldots n$. Define

$$\begin{aligned} a_j &= \sum_{k \in \{k_j\}} E_{kx} B_{kj} + \sum_{l \in \{l_i\}} E_{lx} B_{jl} \\ b_j &= \sum_{k \in \{k_j\}} E_{ky} B_{kj} + \sum_{l \in \{l_i\}} E_{ly} B_{jl} \\ c_j &= \sum_{k \in \{k_j\}} B_{kj} + \sum_{l \in \{l_i\}} B_{jl} > 0. \end{aligned} \tag{21}$$

Observe that c_j is constant and that a_j and b_j are independent of E_{jx} and E_{jy}. Then (16a) and (16b) become

$$-c_j E_{jx}^2 + a_j E_{jx} + (b_j E_{jy} - c_j E_{jy}^2 + Q_j) = 0 \tag{22a}$$

$$a_j E_{jy} - b_j E_{jx} - P_j = 0. \tag{22b}$$

Now define

$$\begin{aligned} A_j &= c_j\left(1 + \frac{b_j^2}{a_j^2}\right) > 0 \\ B_j &= a_j + \frac{b_j^2}{a_j} - \frac{2c_j b_j P_j}{a_j^2} \\ C_j &= -\frac{b_j}{a_j} P_j + \frac{c_j P_j^2}{a_j^2} - Q_j. \end{aligned} \tag{23}$$

A_j, B_j, C_j are independent of E_{jx} and E_{jy}.

Then 22(a) becomes

$$A_j E_{jx}^2 - B_j E_{jx} + C_j = 0 \tag{22a'}$$

so that

$$E_{jx} = \frac{B_j \pm [B_j^2 - 4A_j C_j]^{1/2}}{2A_j} \tag{24a}$$

$$E_{jy} = \frac{b_j E_{jx} + P_j}{a_j}. \tag{24b}$$

In (24b) E_{jx} is computed from (24a). The right hand sides of (24a) and (24b) are independent of E_{jx} and E_{jy}, and are uniquely defined once the algebraic sign in (27a) is determined. The algebraic sign is determined by using the physical constraint that $E_{jx} > 0$, at the flat start $\hat{E}^{FS}$. At $\mathbf{x} = \hat{E}^{FS}$ we have

$$a_j = c_j = \sum_{k\in\{k_j\}} B_{kj} + \sum_{l\in\{l_j\}} B_{jl}; \qquad b_j = 0$$

$$A_j = B_j = c_j > 0$$

$$C_j = \frac{P_j^2}{c_j} - Q_j$$

$$E_{jx} = \frac{c_j \pm [c_j^2 - 4(P_j^2 - c_jQ_j)]^{1/2}}{2c_j} \tag{25a}$$

$$E_{jy} = \frac{P_j}{a_j}. \tag{25b}$$

We consider only the case when impedance and injections satisfy

$$c_j^2 - 4(P_j^2 - c_j Q_j) \geq 0. \tag{26}$$

In order for $E_{jx} > 0$ to hold, the plus sign must be chosen in (25a) and it is therefore adopted in (24a) for *all* $\mathbf{x}$. This defines $\mathbf{t}_k^* = [t_{k1}, t_{k2}]^T = [E_{kx}, E_{ky}]^T$ uniquely. It is assumed that $t_k^* \in Q$ so that condition (3) is fulfilled.

Under the assumptions that have been made, the block Jacobi and block Gauss-Seidel iterations are defined and they converge for *any* $\mathbf{x}^0 \in Q$. We take $\mathbf{x}^0$ to be $\hat{\mathbf{E}}^{FS}$.

6. NUMERICAL EXAMPLES

The use of block Jacobi and block Gauss-Seidel iterations to solve (16) is intended for large systems (n of the order of 500–1500). The load flows in typical practical cases, taken from a collection of IEEE studied systems [12] have been solved using the iterative scheme described here. In these cases n is of the order of 150. The results, which corroborate the theory presented here, are reported in [5]. In order to illustrate the ideas presented here, however, we just consider the small, three bus system shown in Figs. 1 and 2 with the following parameters:

6.1 *An Illustrative three-bus system*

Impedances	*Admittances*	*Injections*
$Z_{12} = .1\sqrt{-1}$	$Y_{12} = -10\sqrt{-1}$	$S_1 = -.5 + .023\sqrt{-1}$
$Z_{13} = .1\sqrt{-1}$	$Y_{13} = -10\sqrt{-1}$	$S_2 = 1.5 + .057\sqrt{-1}$
$Z_{23} = .1\sqrt{-1}$	$Y_{23} = -10\sqrt{-1}$	$S_3 = -1.0 + .036\sqrt{-1}$

Node 1 is taken as the slack bus with $\hat{E}_1 = 1 + 0\sqrt{-1}$. The unknown complex voltages are $\hat{E}_2$, $\hat{E}_3$.

The sets $\{k_j\}$ and $\{l_j\}$ are:

$$\{k_2\} = \{1\} \qquad \{k_3\} = \{1, 2\}$$

$$\{l_2\} = \{3\} \qquad \{l_3\} = \emptyset$$

From (21) we obtain

$$a_2 = 10E_{3x} + 10 \qquad a_3 = 10E_{2x} + 10$$
$$b_2 = 10E_{3y} \qquad b_3 = 10E_{2y}$$
$$c_2 = 20 \qquad c_3 = 20$$

Equation (16a), (16b) can now be written in the form of (25a) and (25b)

$$f_2^{(1)} = -20E_{2x}^2 + (10E_{3x} + 10)E_{2x} + 10E_{3y}E_{2y} - 20E_{2y}^2 + .057 = 0$$
$$f_2^{(2)} = 10E_{3x}E_{2y} - 10E_{3y}E_{2x} - 1.5 = 0$$
$$f_3^{(1)} = -20E_{3x}^2 + (10E_{2x} + 10)E_{3x} + 10E_{2y}E_{3y} - 20E_{3y}^2 + .036 = 0$$
$$f_3^{(2)} = (10E_{2x} + 10)E_{3y} - 10E_{2y}E_{3x} + 1.0 = 0$$

The matrices A_{jk} which are different from zero are:

$$A_{11} = \begin{bmatrix} (10 - 40E_{2x} + 10E_{3x}) & (-40E_{2y} + 10E_{3y}) \\ -10E_{3y} & 10 + 10E_{3x} \end{bmatrix}$$
$$A_{12} = \begin{bmatrix} 10E_{2x} & 10E_{2y} \\ 10E_{2y} & -10E_{2x} \end{bmatrix}$$
$$A_{22} = \begin{bmatrix} (10 - 40E_{3x} + 10E_{2x}) & (-40E_{3y} + 10E_{2y}) \\ -10E_{2y} & 10 + 10E_{2x} \end{bmatrix}$$
$$A_{21} = \begin{bmatrix} 10E_{3x} & 10E_{3y} \\ 10E_{3y} & -10E_{3x} \end{bmatrix}.$$

Both nodes 2 and 3 are connected to the slack node 1. Therefore at $\hat{E}^{FS}$ there is strict block diagonal dominance in row blocks 2 and 3: that is at $\hat{E}^{FS}$ we have

$$\| A_{11}^{-1} \|_\infty \| A_{12} \|_\infty = \frac{1}{20} \cdot 10 = .5 < 1$$
$$\| A_{12}^{-1} \|_\infty \| A_{21} \|_\infty = \frac{1}{20} \cdot 10 = .5 < 1$$

The following table shows the assumption (26) is satisfied

j	$c_j^2 - 4(P_j - c_j Q_j)$
2	387.56
3	484.04

Underrelaxed ($0 < \omega \leq 1$) as well as overrelaxed ($\omega > 1$) block Gauss-Seidel iterations were tested for various values of ω. $\mathbf{x}^0$ was taken to be E^{FS}. The iteration was terminated when $|\mathbf{x}^{p+1} - \mathbf{x}^p| \leq 10^{-d}$. All iterations except $\omega = 1.3$ converged to the solution $\hat{E}'$:

$$\hat{E}'_1 = 1 + 0\sqrt{-1}$$

$$\hat{E}'_2 = 0.99776 + 0.06671\sqrt{-1}$$

$$\hat{E}'_3 = 0.99985 - 0.01671\sqrt{-1}$$

For $\omega = 1.3$, the iteration failed to converge.

The following tables show the number of iterations as a function of ω, for each $d = 3$, 4, 5. On the second line of each table the number of Newton-Raphson (NR) iterations starting at $\mathbf{x}^0 = \hat{E}^{FS}$ is shown

$d = 3$

ω	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0	1.1	1.2
Number of BGS Iterations	17	12	9	8	7	6	6	5	5	7	10	20

Number of NR Iterations = 2

$d = 4$

ω	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0	1.1	1.2
Number of BGS Iterations	50	31	23	18	15	13	11	10	9	10	16	30

Number of NR Iterations = 3

$d = 5$

ω	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0	1.1	1.2
Number of BGS Iterations	95	53	37	28	23	19	16	14	12	19	21	41

Number of NR Iterations = 3

Finally, as a heuristic check on the existence of the closed rectangle Q which contains $\hat{E}^{FS}$ and $\hat{E}'$, which maps into itself under the iteration map, and on which $F'(\hat{E})$ is a blockwise Ω diagonally dominant matrix for $\hat{E} \in Q$, $F'(\hat{E}')$ can be calculated. Substituting $\hat{E} = \hat{E}'$ in the expressions for A_{jk} we obtain

$$\| A_{11}^{-1} \|_\infty \| A_{12} \|_\infty = .66 < 1$$

$$\| A_{22}^{-1} \|_\infty \| A_{21} \|_\infty = .65 < 1$$

Therefore at $\hat{E}'$, the matrix A is, in this case, blockwise strictly diagonally dominant.

6.2 *A realistic thirty node system*

Very similar conclusions to the above were drawn for the standard IEEE test AEP system consisting of 30 nodes. Three types of simulations were performed. The system was treated as:

1.) A lossless system consisting of 29 PQ buses and one slack bus.
2.) A lossless system consisting of 2 PV buses and 28 PQ buses.

Some typical results comparing BGS iterations with NR iterations are:

1.) Initial guess = flat start, $d = 3$

ω	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0	1.05	1.1
Number of BGS Iterations	34	33	57	78	102	109	111	110	107	105	104	*

Number of NR Iterations = 3

* failed to converge

2.) Initial guess = flat start, $d = 3$

ω	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0	1.05	1.1
Number of BGS Iterations	36	51	70	108	120	121	119	116	113	109	118	*

Number of NR Iterations = 3

* failed to converge

Notice that the number of iterations in 1.) and 2.) is approximately the same. From 1.) and 2.) it follows that $\omega \approx 0.1$ is ω^{opt}. Using smaller increments (0.1) around $\omega = 0.1$ confirmed that ω^{opt} is indeed approximately 0.1. For this ω we have studied the effects of varying the initial conditions from the flat start, although this is not covered by our theory. The initial voltage was always taken to be real $E_{ix}^{o} = E_{x}^{o}$, real for all i, $E_{y}^{o} = 0$, $d = 4$, $\omega = 0.1$. We also tested the Newton-Raphson method starting at each initial guess. The results are as follows:

E°	.9	.8	.7	.6	.5	.4	.3	.2
Number of BGS Iterations	35	54	86	116	142	162	*	*
Number of NR Iterations	4	4	5	7	*	*	*	*

E_x°	.9	.8	.7	.6	.5	.4	.3	.2
Number of BGS Iterations	56	85	118	189	238	292	*	*
Number of NR Iterations	4	4	5	7	*	*	*	*

* failed to converge

It is clear that when NR converges it requires fewer iterations to achieve the desired accuracy (10^{-3}) than BGS. The cost per iteration for NR is, of course, much greater. On the other hand, BGS converges for initial guesses that are considerably further from the flat start than NR.

When making the sample runs, it was observed that the nonconverging cases were discovered much faster in the BGS method than in the NR method. In the BGS method the nonconverging cases are determined easily via the nonexistence of of a real solution in (24a), because of a negative value under the square root. In using the NR method one has to actually allow a large number of NR steps before concluding that the iteration diverges.

REFERENCES

1. H. Dommel and W. Tinney, Optimal power flow solutions, *IEEE Trans. on Power Apparatus and Systems*, 1866–1876 (1968).
2. O. I. Elgerd, *Electric Energy Systems Theory: An Introduction*, McGraw-Hill, New York (1971).
3. David G. Feingold and Richard S. Varga, Block Diagonally Dominant Matrices and Generalizations of the Gerschgorin Circle Theorem, *Pacific J. Math.*, **12**, 1241–1250 (1962).
4. F. D. Galiana, Analytic Properties of the Load Flow Problem, IEEE International Symposium on Circuits and Systems (1977).
5. M. Ilić-Spong and J. Zaborszky, A Different Approach to Load Flow, PICA Conference, Philadelphia, May 1981, *IEEE Trans. on Power Apparatus and Systems*, 168–179 (Jan. 1982).
6. Marija Ilić-Spong, A New Approach to Load Flow Studies of Electric Power Systems Using Complex Power-Complex Voltage Graphs, Doctoral Dissertation, Department of Systems Science and Mathematics, Washington University, St. Louis, Missouri (1980).
7. Jorge J. Moré, Nonlinear Generalizations of Matrix Diagonal Dominance With Application to Gauss-Seidel Iterations, *SIAM J. Numer. Anal.*, **9**, No. 2, 357–378 (1972).
8. H. S. Price, A Practical Application of Block Diagonally Dominant Matrices, *Math. of Comp.*, **19**, 307–313 (1965).
9. G. W. Stagg and A. H. El-Abiad, *Computer Methods in Power Systems Analysis*, McGraw-Hill, New York (1968).
10. Brian Stott, Reveiw of Load-Flow Calculation Methods, *Proc. IEEE*, **62**, No. 7, 916–929 (1974).
11. W. Tinny and C. Hart, Power Flow Solution by Newton's Method, *IEEE Trans. on Power Apparatus and Systems*, 1449–1460 (1967).
12. J. Zaborszky, K. W. Whang, and K. V. Prasad, Operation of the large interconnected power system by decision and control in emergencies, Conference on Systems Engineering for Power, Davos, Switzerland, (1979).