Electronic Transactions on Numerical Analysis.
Volume 58, pp. 621–628, 2023.

ISSN 1068–9613.
DOI: 10.1553/etna_vol58s621

ETNA
Kent State University and
Johann Radon Institute (RICAM)

A GENERALIZED SASSENFELD CRITERION AND ITS RELATION TO H-MATRICES*

THOMAS P. WIHLER†

Abstract. The starting point of this note is a decades-old yet little-noticed sufficient condition, presented by Sassenfeld in 1951, for the convergence of the classical Gauß–Seidel method. The purpose of the present paper is to shed new light on Sassenfeld's criterion and to demonstrate that it is closely related to H-matrices. In particular, our main result yields a novel characterization of H-matrices. In addition, a new convergence estimate for iterative linear solvers, which involve H-matrix preconditioners, is briefly discussed.

Key words. Sassenfeld criterion, convergence of iterative linear solvers, splitting methods, Gauß–Seidel scheme, preconditioning, H-matrices

AMS subject classifications. 15B48, 65F08, 65F10

1. Introduction. The Gauß–Seidel method is amongst the most classical numerical schemes for the iterative solution of systems of linear equations. Traditionally, in many numerical analysis textbooks, convergence is established for matrices that are either strictly diagonally dominant or symmetric positive definite. Only a few authors (see, e.g., [13, Theorem 4.16]) point to a less-standard convergence criterion for the Gauß–Seidel scheme that was introduced by Sassenfeld in his paper [10]: Given a matrix $\mathbf{A} = [a_{ij}] \in \mathbb{C}^{m\times m}$ with non-vanishing diagonal entries, i.e., $a_{ii} \neq 0$, for each $i = 1, \ldots, m$, define non-negative real numbers $s_1, \ldots, s_m$ iteratively by

$$s_i = \frac{1}{|a_{ii}|}\left(\sum_{j<i} |a_{ij}|s_j + \sum_{j>i} |a_{ij}|\right), \qquad i = 1, \ldots, m. \tag{1.1}$$

Sassenfeld has proved that the condition

$$0 \leqslant s_i < 1, \qquad \forall i = 1, \ldots, m, \tag{1.2}$$

is sufficient for the convergence of the Gauß–Seidel iteration. Matrices that satisfy this property (which is closely related to generalized diagonal dominance; see, e.g., [6]) were discussed recently in [1].

The purpose of the present note is to show that there is a more general principle behind Sassenfeld's original work which is intimately related to H-matrices. To illustrate this observation, we note that (1.1) can be written in matrix form as

$$(|\mathbf{D}| - |\mathbf{L}|)\mathbf{s} = |\mathbf{U}|\mathbf{e}, \tag{1.3}$$

where the matrix $\mathbf{A} = \mathbf{L} + \mathbf{D} + \mathbf{U}$ is decomposed in the usual way into the (strict) lower and upper triangular parts $\mathbf{L} = \mathrm{tril}(\mathbf{A})$ and $\mathbf{U} = \mathrm{triu}(\mathbf{A})$, respectively, and the diagonal part $\mathbf{D} = \mathrm{diag}(\mathbf{A})$. Furthermore, $|[\star]|$ signifies the modulus of a matrix $[\star]$ taken entry-wise, $\mathbf{s} = (s_1, \ldots, s_m)$ is a vector that contains the iteratively defined non-negative real numbers $s_1, \ldots, s_m$ from (1.1), and

$$\mathbf{e} = (1, \ldots, 1)^\intercal \in \mathbb{R}^m \tag{1.4}$$

*Received January 17, 2023. Accepted June, 26, 2023. Published online on December 7, 2023. Recommended by Froilan Dopico.
†Mathematics Institute, University of Bern, CH-3012 Bern, Switzerland (thomas.wihler@unibe.ch).

ETNA
Kent State University and
Johann Radon Institute (RICAM)

is the (column) vector with all components 1. More generally, for appropriate matrices $\mathbf{P} \in \mathbb{C}^{m\times m}$, we consider the splitting

$$\mathbf{A} = \operatorname{off}(\mathbf{P}) + \operatorname{diag}(\mathbf{P}) + (\mathbf{A} - \mathbf{P}),$$

where $\operatorname{off}([\star])$ denotes the off-diagonal part of a matrix $[\star]$. Then, define the vector $\mathbf{s} \in \mathbb{R}^m$ to be the solution (if it exists) of the system

$$(|\operatorname{diag}(\mathbf{P})| - |\operatorname{off}(\mathbf{P})|)\mathbf{s} = |\mathbf{A} - \mathbf{P}|\mathbf{e}. \tag{1.5}$$

For instance, in the context of the Gauß–Seidel scheme, letting $\mathbf{P} := \mathbf{L} + \mathbf{D}$, with $\mathbf{L}$ and $\mathbf{D}$ as above, we notice that (1.5) translates immediately into (1.3). In this work, we will focus on matrices $\mathbf{A}$ and $\mathbf{P}$ for which the components of the solution vector $\mathbf{s}$ of the linear system (1.5) satisfy the Sassenfeld criterion (1.2).

Outline. We begin our work by reviewing the class of H-matrices (see Section 2), which was originally introduced in [8] and which plays a crucial role in the convergence of iterative splitting methods (especially, the Jacobi, Gauß–Seidel, and SOR schemes); in the context of this paper, such matrices are exactly those for which the system (1.5) is non-singular. In Section 3 we continue by introducing the so-called *Sassenfeld index*, which is an essential quantity for our analysis, and derive some basic estimates. Subsequently, in Section 4, based on the previously defined Sassenfeld index, we will focus on all matrices for which the bounds (1.2) for the solution vector $\mathbf{s}$ of (1.5) can be achieved; such matrices will be said to satisfy the *generalized Sassenfeld criterion*. Our main result (Theorem 4.5) will show that any matrix in $\mathbb{C}^{m\times m}$ is a (non-singular) H-matrix *if and only if* it fulfills the generalized Sassenfeld criterion; in this regard, our work provides *a new characterization of H-matrices*. In addition, a computational verification procedure is proposed (see Proposition 3.4); cf. the related papers [4, 7]. Finally, we conclude this article with a few remarks in Section 5.

Notation. For any vectors or matrices $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{m\times n}$, we use the notation $\mathbf{X} \succeq \mathbf{Y}$ (or $\mathbf{X} \succ \mathbf{Y}$) to indicate that all entries of the difference $\mathbf{X} - \mathbf{Y} \in \mathbb{R}^{m\times n}$ are non-negative (respectively positive). Furthermore, for a matrix $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{m\times n}$, we denote by $\|\mathbf{A}\|_\infty := \max_{1\leqslant i\leqslant m} \sum_{j=1}^n |a_{ij}|$ the standard ∞-norm. Moreover, we signify by $\varrho(\mathbf{A})$ the spectral radius of a square matrix $\mathbf{A} \in \mathbb{C}^{m\times m}$, and $\mathbf{I}_m \in \mathbb{C}^{m\times m}$ is the identity matrix.

2. A brief review of H-matrices.

We will denote by $\mathcal{H}_m$ the subset of all H-matrices in $\mathbb{C}^{m\times m}$. This set was originally introduced in [8] (see also [2, 3, 12]), and it consists of all matrices $\mathbf{A} = [a_{ij}] \in \mathbb{C}^{m\times m}$ for which the associated comparison matrix, given by

$$\mathfrak{M}(\mathbf{A}) := |\operatorname{diag}(\mathbf{A})| - |\operatorname{off}(\mathbf{A})| = \begin{cases} -|a_{ij}|, & \text{if } i \neq j, \\ +|a_{ii}|, & \text{if } i = j, \end{cases} \qquad 1 \leqslant i, j \leqslant m,$$

is a non-singular M-matrix, i.e., it takes the form $\mathfrak{M}(\mathbf{A}) = r\mathbf{I}_m - \mathbf{B}$, for a matrix $\mathbf{B} \succeq \mathbf{0}$, with $r > \varrho(\mathbf{B})$.

We collect a few well-known facts about H-matrices that are instrumental for the present work.

(F1) We first remark that matrices in $\mathcal{H}_m$ are non-singular; see [8].

(F2) Moreover, it is well known (see, e.g., [5, Theorem 5']) that $\mathbf{A} \in \mathcal{H}_m$ if and only if there is a positive real vector $\mathbf{u} \succ \mathbf{0}$ such that $\mathfrak{M}(\mathbf{A})\mathbf{u} \succ \mathbf{0}$; in individual components, this means that there are positive numbers $u_1, \ldots, u_m > 0$ such that

$$|a_{ii}|u_i > \sum_{j\neq i} |a_{ij}|u_j, \qquad \forall i = 1, \ldots, m.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Incidentally, this property refers to the notion of generalized diagonal dominance (by rows); cf., e.g., [6]. In particular, the above bound implies that the diagonal entries of any matrix in $\mathcal{H}_m$ are all non-zero.

(F3) Furthermore, for $\mathbf{A} \in \mathcal{H}_m$, since $\mathfrak{M}(\mathbf{A})$ is a non-singular M-matrix, it follows that $\mathfrak{M}(\mathbf{A})^{-1} \succeq \mathbf{0}$; see, e.g., [9].

(F4) Finally, for any matrix $\mathbf{A} \in \mathcal{H}_m$, it holds that

$$\varrho\left(|\mathrm{diag}(\mathbf{A})|^{-1}|\mathrm{off}(\mathbf{A})|\right) = \varrho\left(\mathrm{diag}(\mathfrak{M}(\mathbf{A}))^{-1}\,\mathrm{off}(\mathfrak{M}(\mathbf{A})\right) < 1;$$

see, e.g., [11, Theorem 1 (vii)].

3. Sassenfeld index. For a non-singular matrix $\mathbf{A} \in \mathbb{C}^{m\times m}$, a right-hand side vector $\mathbf{b} \in \mathbb{C}^m$, and an arbitrary starting vector $\mathbf{x}^{(0)} \in \mathbb{C}^m$, we will be interested in the iterative splitting scheme

$$\mathbf{P}\mathbf{x}^{(k+1)} = (\mathbf{P}-\mathbf{A})\mathbf{x}^{(k)} + \mathbf{b}, \qquad k \geqslant 0, \tag{3.1}$$

for the solution of the linear system

$$\mathbf{A}\mathbf{x} = \mathbf{b}. \tag{3.2}$$

The focus of this work will be on preconditioners $\mathbf{P} \in \mathcal{H}_m$.

From fact (F3) above, for $\mathbf{P} \in \mathcal{H}_m$, we infer that the vector defined by

$$\mathbf{s}(\mathbf{A},\mathbf{P}) := \mathfrak{M}(\mathbf{P})^{-1}|\mathbf{A}-\mathbf{P}|\mathbf{e} \succeq \mathbf{0}, \tag{3.3}$$

with $\mathbf{e} \in \mathbb{R}^m$ from (1.4), is well defined and contains only non-negative components.

DEFINITION 3.1 (Sassenfeld index). *The* Sassenfeld index *of a matrix* $\mathbf{A} \in \mathbb{C}^{m\times m}$ *with respect to a preconditioner* $\mathbf{P} \in \mathcal{H}_m$ *is defined by* $\mu(\mathbf{A},\mathbf{P}) := \|\mathbf{s}(\mathbf{A},\mathbf{P})\|_\infty$, *with the vector* $\mathbf{s}(\mathbf{A},\mathbf{P})$ *from* (3.3).

The essence of the Sassenfeld index defined above is that it allows to control the norm $\left\|\mathbf{I}_m - \mathbf{P}^{-1}\mathbf{A}\right\|_\infty$ of the iteration matrix in the splitting method (3.1) in a non-standard way.

PROPOSITION 3.2. *Let* $\mathbf{A} \in \mathbb{C}^{m\times m}$ *be a non-singular matrix and* $\mathbf{P} \in \mathcal{H}_m$. *Then it holds that*

$$\left\|\mathbf{I}_m - \mathbf{P}^{-1}\mathbf{A}\right\|_\infty \leqslant \mu(\mathbf{A},\mathbf{P}). \tag{3.4}$$

Proof. Consider an arbitrary vector $\mathbf{y} \in \mathbb{C}^m$ with $\|\mathbf{y}\|_\infty = 1$. Defining $\mathbf{R} = \mathbf{P}-\mathbf{A}$, we let

$$\mathbf{x} := \mathbf{P}^{-1}\mathbf{R}\mathbf{y} = \mathbf{P}^{-1}(\mathbf{P}-\mathbf{A})\mathbf{y} = (\mathbf{I}_m - \mathbf{P}^{-1}\mathbf{A})\mathbf{y}. \tag{3.5}$$

Note first that $\mathrm{diag}(\mathbf{P})\mathbf{x} + \mathrm{off}(\mathbf{P})\mathbf{x} = \mathbf{R}\mathbf{y}$. Taking moduli results in

$$\mathfrak{M}(\mathbf{P})|\mathbf{x}| \preceq |\mathbf{R}||\mathbf{y}| \preceq |\mathbf{R}|\mathbf{e}.$$

Recalling that $\mathfrak{M}(\mathbf{P})^{-1} \succeq \mathbf{0}$, cf. fact (F3) above, and employing (3.3), we deduce that

$$|\mathbf{x}| \preceq \mathfrak{M}(\mathbf{P})^{-1}|\mathbf{R}|\mathbf{e} = \mathbf{s}(\mathbf{A},\mathbf{P}).$$

Therefore, using (3.5), we infer that

$$\left\|(\mathbf{I}_m - \mathbf{P}^{-1}\mathbf{A})\mathbf{y}\right\|_\infty = \|\mathbf{x}\|_\infty \leqslant \|\mathbf{s}(\mathbf{A},\mathbf{P})\|_\infty,$$

which yields (3.4). ∎

ETNA
Kent State University and
Johann Radon Institute (RICAM)

COROLLARY 3.3 (Invertibility). *Given a matrix* $\mathbf{A}$ *and a preconditioner* $\mathbf{P} \in \mathcal{H}_m$. *Then the matrix* $\mathbf{A}_\tau = \mathbf{A} + \tau\mathbf{P}$ *is non-singular whenever* $|\tau + 1| > \mu(\mathbf{A}, \mathbf{P})$.

Proof. We apply a contradiction argument. To this end, suppose that there exists $\mathbf{v} \in \mathbb{C}^m$, $\|\mathbf{v}\|_\infty = 1$, such that $\mathbf{A}_\tau \mathbf{v} = \mathbf{0}$. Then it holds that $(\tau + 1)\mathbf{P}\mathbf{v} = (\mathbf{P} - \mathbf{A})\mathbf{v}$, and thus $(\tau + 1)\,\mathbf{v} = \mathbf{P}^{-1}(\mathbf{P} - \mathbf{A})\mathbf{v}$. Taking norms and using (3.4) yields

$$|\tau + 1| = \left\|(\mathbf{I}_m - \mathbf{P}^{-1}\mathbf{A})\mathbf{v}\right\|_\infty \leqslant \left\|\mathbf{I}_m - \mathbf{P}^{-1}\mathbf{A}\right\|_\infty \leqslant \mu(\mathbf{A}, \mathbf{P}).$$

This completes the proof. ∎

We note that the vector $\mathbf{s}(\mathbf{A}, \mathbf{P})$ from (3.3) can be computed approximately by iteration. Indeed, if $\mathbf{P} \in \mathcal{H}_m$, then the diagonal entries of $\mathbf{P}$ do not vanish, cf. fact (F2) above, and the iterative scheme given by

$$|\mathrm{diag}(\mathbf{P})|\mathbf{s}^{(k+1)} = |\mathrm{off}(\mathbf{P})|\mathbf{s}^{(k)} + |\mathbf{A} - \mathbf{P}|\mathbf{e}, \qquad k \geqslant 0, \tag{3.6}$$

converges to the vector $\mathbf{s}(\mathbf{A}, \mathbf{P})$ from (3.3) for any initial vector $\mathbf{s}^{(0)} \in \mathbb{C}^m$ by fact (F4). Furthermore, the following result provides a computational upper bound for the Sassenfeld index.

PROPOSITION 3.4 (Iterative estimation of the Sassenfeld index). *Consider a matrix* $\mathbf{A} \in \mathbb{C}^{m\times m}$ *and a preconditioner* $\mathbf{P} \in \mathcal{H}_m$. *Then, there exists an initial vector* $\mathbf{s}^{(0)} \in \mathbb{R}^m$ *such that*

$$|\mathbf{A} - \mathbf{P}|\mathbf{e} \preceq \mathfrak{M}(\mathbf{P})\mathbf{s}^{(0)}. \tag{3.7}$$

Furthermore, if the iteration (3.6) *is initiated by a vector* $\mathbf{s}^{(0)}$ *(for* $k = 0$*) that satisfies* (3.7), *then it holds that* $\mu(\mathbf{A}, \mathbf{P}) \leqslant \left\|\mathbf{s}^{(k)}\right\|_\infty$, *for all* $k \geqslant 0$, *and* $\lim_{k\to\infty} \left\|\mathbf{s}^{(k)}\right\|_\infty = \mu(\mathbf{A}, \mathbf{P})$.

Proof. The existence of a vector $\mathbf{s}^{(0)}$ that satisfies (3.7) is immediately established upon setting $\mathbf{s}^{(0)} := \mathbf{s}(\mathbf{A}, \mathbf{P})$, cf. (3.3). Now consider any vector $\mathbf{s}^{(0)} \in \mathbb{R}^m$ that fulfills (3.7). Then, from (3.6) with $k = 0$, we have

$$|\mathrm{diag}(\mathbf{P})|\left(\mathbf{s}^{(1)} - \mathbf{s}^{(0)}\right) = -\mathfrak{M}(\mathbf{P})\mathbf{s}^{(0)} + |\mathbf{A} - \mathbf{P}|\mathbf{e} \preceq \mathbf{0},$$

which shows that $\mathbf{s}^{(1)} - \mathbf{s}^{(0)} \preceq \mathbf{0}$. Hence, by induction, since $\mathfrak{M}(\mathbf{P})^{-1} \succeq \mathbf{0}$, cf. fact (F3), we note that

$$\mathbf{s}^{(k+1)} - \mathbf{s}^{(k)} = \left|\mathrm{diag}(\mathbf{P})^{-1}\,\mathrm{off}(\mathbf{P})\right|\left(\mathbf{s}^{(k)} - \mathbf{s}^{(k-1)}\right) \preceq \mathbf{0}, \qquad \forall k \geqslant 1. \tag{3.8}$$

Using that $\varrho\left(\left|\mathrm{diag}(\mathbf{P})^{-1}\,\mathrm{off}(\mathbf{P})\right|\right) < 1$, cf. fact (F4), we infer that the iteration (3.6) converges to $\mathbf{s}(\mathbf{A}, \mathbf{P})$ from (3.3). Moreover, from (3.3) and (3.6) we deduce the identity

$$\begin{aligned}
\mathfrak{M}(\mathbf{P})\mathbf{s}(\mathbf{A}, \mathbf{P}) &= |\mathbf{A} - \mathbf{P}|\mathbf{e} \\
&= |\mathrm{diag}(\mathbf{P})|\mathbf{s}^{(k+1)} - |\mathrm{off}(\mathbf{P})|\mathbf{s}^{(k)} \\
&= \mathfrak{M}(\mathbf{P})\mathbf{s}^{(k+1)} + |\mathrm{off}(\mathbf{P})|\left(\mathbf{s}^{(k+1)} - \mathbf{s}^{(k)}\right),
\end{aligned}$$

for all $k \geqslant 0$. Exploiting again that $\mathfrak{M}(\mathbf{P})^{-1} \succeq \mathbf{0}$ and upon involving (3.8), we arrive at

$$\mathbf{s}(\mathbf{A}, \mathbf{P}) = \mathbf{s}^{(k+1)} + \mathfrak{M}(\mathbf{P})^{-1}|\mathrm{off}(\mathbf{P})|\left(\mathbf{s}^{(k+1)} - \mathbf{s}^{(k)}\right) \preceq \mathbf{s}^{(k+1)} \preceq \mathbf{s}^{(k)}. \tag{3.9}$$

Since $\mathbf{s}(\mathbf{A}, \mathbf{P})$ and $\mathbf{s}^{(k)}$ are both non-negative, the asserted bound follows. ∎

The ensuing result, which immediately follows from (3.9), allows for an estimate of the Sassenfeld index without solving the system (3.3).

COROLLARY 3.5. *Given a matrix* $\mathbf{A} \in \mathbb{C}^{m\times m}$. *Furthermore, let* $\mathbf{P} \in \mathcal{H}_m$*, and let* $\mathbf{v} \in \mathbb{R}^m$ *be a non-negative vector such that*

$$|\mathbf{A} - \mathbf{P}|\mathbf{e} \preceq \mathfrak{M}(\mathbf{P})\mathbf{v}. \tag{3.10}$$

Then, it holds that $\mu(\mathbf{A},\mathbf{P}) \leqslant \|\mathbf{v}\|_\infty$.

EXAMPLE 3.6 (Jacobi preconditioner). If $\mathbf{P} = \operatorname{diag}(\mathbf{A})$ is non-singular, then the bound (3.10) is fulfilled for any vector $\mathbf{v}$ with components

$$v_i \geqslant \frac{1}{|a_{ii}|}\sum_{j\neq i}|a_{ij}|, \qquad 1 \leqslant i \leqslant m, \tag{3.11}$$

and we have $\mu(\mathbf{A}, \operatorname{diag}(\mathbf{A})) \leqslant \max_{1\leqslant i\leqslant m} v_i$. For instance, for the classical finite difference matrix

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & & \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix} \in \mathbb{R}^{m\times m}, \tag{3.12}$$

with $m \geqslant 3$, and the vector $\mathbf{v} = (\frac{1}{2}, 1, \dots, 1, \frac{1}{2})^\mathsf{T}$, which satisfies (3.11) with equality, we have $\max_{1\leqslant i\leqslant m} v_i = 1 = \mu(\mathbf{A}, 2\mathbf{I}_m)$.

EXAMPLE 3.7 (Gauß–Seidel preconditioner). If $\mathbf{P} = \operatorname{tril}(\mathbf{A}) + \operatorname{diag}(\mathbf{A})$ is non-singular, then (3.10) translates into

$$\sum_{j>i}|a_{ij}| \leqslant |a_{ii}|v_i - \sum_{j<i}|a_{ij}|v_j, \qquad 1 \leqslant i \leqslant m, \tag{3.13}$$

which is essentially the recursive relation (1.1) for Sassenfeld's original criterion. For the matrix $\mathbf{A}$ from (3.12), for $m \geqslant 2$, and

$$\mathbf{v} = \left(1 - 2^{-1}, 1 - 2^{-2}, \dots, 1 - 2^{1-m}, \tfrac{1}{2} - 2^{-m}\right)^\mathsf{T},$$

for which equality holds in (3.13), it is elementary to verify that

$$\max_{1\leqslant i\leqslant m} v_i = 1 - 2^{1-m} = \mu(\mathbf{A}, \operatorname{diag}(\mathbf{A}) + \operatorname{tril}(\mathbf{A}));$$

this yields (1.2).

4. A characterization of H-matrices by the Sassenfeld index.

4.1. Generalized Sassenfeld criterion. We are now ready to disclose a connection between the Sassenfeld index and H-matrices. Our definition of a generalized Sassenfeld criterion (see Definition 4.1 below) is motivated by the work [1], where the special case of all matrices $\mathbf{A} \in \mathbb{C}^{m\times m}$ with $\mu(\mathbf{A},\mathbf{P}) < 1$, with $\mathbf{P} = \operatorname{tril}(\mathbf{A}) + \operatorname{diag}(\mathbf{A})$ being the Gauß–Seidel preconditioner, has been discussed. In this specific situation, the system (3.3) takes the (lower-triangular) form

$$|\operatorname{diag}(\mathbf{A})|\mathbf{s} = |\operatorname{tril}(\mathbf{A})|\mathbf{s} + |\operatorname{triu}(\mathbf{A})|\mathbf{e},$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

which is a simple forward solve for $\mathbf{s}$. Convergence of the Gauß–Seidel method is guaranteed if $\|\mathbf{s}\|_\infty < 1$; this is the key observation in Sassenfeld's original work [10].

More generally, for preconditioners $\mathbf{P} \in \mathcal{H}_m$ in the current paper, we propose the following definition:

DEFINITION 4.1 (Generalized Sassenfeld criterion). *A matrix* $\mathbf{A} \in \mathbb{C}^{m\times m}$ *is said to satisfy the* generalized Sassenfeld criterion *if there exists a preconditioner* $\mathbf{P} \in \mathcal{H}_m$ *such that* $\mu(\mathbf{A},\mathbf{P}) < 1$.

REMARK 4.2. From Corollary 3.3, for $\tau = 0$, we immediately deduce that every matrix which fulfills the generalized Sassenfeld criterion is non-singular.

REMARK 4.3. The verification of the generalized Sassenfeld criterion requires the existence of a suitable preconditioning matrix $\mathbf{P} \in \mathcal{H}_m$ such that $\mu(\mathbf{A},\mathbf{P}) < 1$. Hence, from Proposition 3.2, we need $\mathbf{P}^{-1}\mathbf{A} \approx \mathbf{I}_m$ in the sense that $\left\|\mathbf{I}_m - \mathbf{P}^{-1}\mathbf{A}\right\|_\infty \leqslant \mu(\mathbf{A},\mathbf{P}) < 1$.

The following proposition provides a condition number estimate for the preconditioned matrix $\mathbf{P}^{-1}\mathbf{A}$ in terms of the Sassenfeld index:

PROPOSITION 4.4 (Condition number bound). *Suppose that* $\mathbf{A}$ *satisfies the generalized Sassenfeld criterion for a suitable preconditioner* $\mathbf{P} \in \mathcal{H}_m$ *with* $\mu(\mathbf{A},\mathbf{P}) < 1$. *Then, for the condition number (with respect to the* ∞*-norm), the bound*

$$\kappa_\infty(\mathbf{P}^{-1}\mathbf{A}) \leqslant \frac{1+\mu(\mathbf{A},\mathbf{P})}{1-\mu(\mathbf{A},\mathbf{P})}$$

holds true.

Proof. Let $\mathbf{C} := \mathbf{P}^{-1}\mathbf{A}$. From Proposition 3.2, we deduce the bound

$$\|\mathbf{C}\|_\infty \leqslant 1 + \left\|\mathbf{I}_m - \mathbf{P}^{-1}\mathbf{A}\right\|_\infty \leqslant 1 + \mu(\mathbf{A},\mathbf{P}).$$

Moreover, applying a Neumann series, we deduce the estimate

$$\left\|\mathbf{C}^{-1}\right\|_\infty = \left\|(\mathbf{I}_m - (\mathbf{I}_m - \mathbf{C}))^{-1}\right\|_\infty \leqslant \frac{1}{1-\|\mathbf{I}_m - \mathbf{C}\|_\infty} \leqslant \frac{1}{1-\mu(\mathbf{A},\mathbf{P})}.$$

This concludes the proof. ∎

4.2. The generalized Sassenfeld criterion and H-matrices. We will now establish the main result of this paper.

THEOREM 4.5. *For any* $m \geqslant 1$, *a matrix* $\mathbf{A} \in \mathbb{C}^{m\times m}$ *is a (non-singular) H-matrix if and only if it satisfies the generalized Sassenfeld criterion.*

Proof. If $\mathbf{A} \in \mathcal{H}_m$, then $\mu(\mathbf{A},\mathbf{A}) = 0$, i.e., $\mathbf{A}$ satisfies the generalized Sassenfeld criterion. Conversely, for $\mathbf{A} \in \mathbb{C}^{m\times m}$, suppose that there exists $\mathbf{P} \in \mathcal{H}_m$ with $\mu(\mathbf{A},\mathbf{P}) < 1$. Then, writing (3.3) component-wise, there are non-negative real numbers $0 \leqslant s_i < 1$, $i = 1,\ldots,m$, such that

$$|p_{ii}|s_i - \sum_{j\neq i}|p_{ij}|s_j = \sum_{j=1}^m |a_{ij} - p_{ij}|, \qquad i = 1,\ldots,m.$$

Letting

$$\delta_i := \frac{1}{|a_{ii}|}\sum_{j=1}^m (1-s_j)|a_{ij}-p_{ij}| \geqslant 0, \qquad i = 1,\ldots,m, \tag{4.1}$$

and rearranging terms, we observe the identity

$$\delta_i|a_{ii}| + \sum_{j\neq i}\left(|p_{ij}| + |a_{ij}-p_{ij}|\right)s_j = \left(|p_{ii}| - |a_{ii}-p_{ii}|\right)s_i,$$

for each $i = 1, \ldots, m$. Applying the triangle inequality on either side, it follows that

$$\sum_{j \neq i} |a_{ij}| s_j \leqslant |a_{ii}|(s_i - \delta_i), \qquad i = 1, \ldots, m. \tag{4.2}$$

Furthermore, recalling fact (F2), there are positive numbers $u_1, \ldots, u_m > 0$ such that

$$\sum_{j \neq i} |p_{ij}| u_j < |p_{ii}| u_i, \tag{4.3}$$

for each $i = 1, \ldots, m$. Introduce positive numbers $\vartheta_i := \alpha s_i + u_i > 0$, $i = 1, \ldots, m$, where $\alpha \geqslant 0$ will be specified later; see (4.5) below. Then, for $1 \leqslant i \leqslant m$, we have

$$\sum_{j \neq i} |a_{ij}| \vartheta_j = \alpha \sum_{j \neq i} |a_{ij}| s_j + \sum_{j \neq i} |p_{ij}| u_j + \sum_{j \neq i} \left(|a_{ij}| - |p_{ij}|\right) u_j.$$

Employing (4.2) and (4.3), for each $i = 1, \ldots, m$, we derive the estimate

$$\sum_{j \neq i} |a_{ij}| \vartheta_j < \alpha(s_i - \delta_i)|a_{ii}| + |p_{ii}| u_i + \sum_{j \neq i} \left(|a_{ij}| - |p_{ij}|\right) u_j.$$

Thus, we obtain

$$\sum_{j \neq i} |a_{ij}| \vartheta_j < |a_{ii}| \vartheta_i - \alpha \delta_i |a_{ii}| + u_i \left(|p_{ii}| - |a_{ii}|\right) + \sum_{j \neq i} \left(|a_{ij}| - |p_{ij}|\right) u_j, \tag{4.4}$$

for each $i = 1, \ldots, m$. Now choose $\alpha \geqslant 0$ sufficiently large so that

$$\alpha \delta_i |a_{ii}| \geqslant u_i \left(|p_{ii}| - |a_{ii}|\right) + \sum_{j \neq i} \left(|a_{ij}| - |p_{ij}|\right) u_j, \qquad \forall i \in \mathcal{I}, \tag{4.5}$$

where $\mathcal{I}$ indicates the set of all indices $1 \leqslant i \leqslant m$ for which $\delta_i > 0$ in (4.1); we let $\alpha = 0$ if $\mathcal{I} = \emptyset$ is empty. Now we distinguish two separate cases:

(i) If $\delta_i = 0$, then exploiting that $0 \leqslant s_j < 1$, for each $j = 1, \ldots, m$, we notice from (4.1) that $a_{ij} = p_{ij}$, for all $j = 1, \ldots, m$. Then, from (4.4), we infer that $\sum_{j \neq i} |a_{ij}| \vartheta_j < |a_{ii}| \vartheta_i$, for all $i \notin \mathcal{I}$.

(ii) Otherwise, if $\delta_i > 0$, then recalling α from (4.5), we obtain that $\sum_{j \neq i} |a_{ij}| \vartheta_j < |a_{ii}| \vartheta_i$, for all $i \in \mathcal{I}$.

In summary, we conclude that $\sum_{j \neq i} |a_{ij}| \vartheta_j < |a_{ii}| \vartheta_i$, for each $i = 1, \ldots, m$, which implies that $\mathbf{A} \in \mathcal{H}_m$; cf. fact (F2). ☐

4.3. Application to splitting methods. In the context of linear solvers, the following generalization of Sassenfeld's result [10] for the Gauß–Seidel scheme is an immediate consequence of Proposition 3.2 and Theorem 4.5.

PROPOSITION 4.6 (Iterative solvers). *For $\mathbf{A} \in \mathcal{H}_m$ and any given vector $\mathbf{b} \in \mathbb{C}^m$, consider the linear system (3.2). Then, for any preconditioner $\mathbf{P} \in \mathcal{H}_m$ with $\mu(\mathbf{A}, \mathbf{P}) < 1$, the iteration (3.1) converges to the unique solution of (3.2) for every starting vector $\mathbf{x}^{(0)} \in \mathbb{C}^m$. Furthermore, the following* a priori *error bound holds:*

$$\left\| \mathbf{x} - \mathbf{x}^{(k)} \right\|_\infty \leqslant \mu(\mathbf{A}, \mathbf{P})^k \left\| \mathbf{x} - \mathbf{x}^{(0)} \right\|_\infty,$$

for any $k \geqslant 0$.

5. Conclusions. Inspired by Sassenfeld's historical convergence criterion for the classical Gauß–Seidel scheme, we have introduced the notion of the *Sassenfeld index* (with respect to H-matrix preconditioners), which, in turn, gives rise to a *generalized Sassenfeld criterion* discussed in this work. Our main result shows that a matrix is a (non-singular) H-matrix if and only if it satisfies the generalized Sassenfeld criterion, thereby yielding a new characterization for such matrices. Moreover, an iterative procedure for the computational verification of the proposed generalized Sassenfeld criterion is provided.

REFERENCES

[1] R. BAUMANN AND T. P. WIHLER, *A note on Sassenfeld matrices*, Elem. Math., 72 (2017), pp. 62–65.
[2] A. BERMAN AND R. J. PLEMMONS, *Nonnegative Matrices in the Mathematical Sciences*, SIAM, Philadelphia, 1994.
[3] R. BRU, C. CORRAL, I. GIMENEZ, AND J. MAS, *Classes of general H-matrices*, Linear Algebra Appl., 429 (2008), pp. 2358–2366.
[4] R. BRU, I. GIMÉNEZ, AND A. HADJIDIMOS, *Is $A \in \mathbb{C}^{n,n}$ a general H-matrix?*, Linear Algebra Appl., 436 (2012), pp. 364–380.
[5] K. FAN, *Topological proofs for certain theorems on matrices with non-negative elements*, Monatsh. Math., 62 (1958), pp. 219–237.
[6] K. R. JAMES AND W. RIHA, *Convergence criteria for successive overrelaxation*, SIAM J. Numer. Anal., 12 (1975), pp. 137–143.
[7] K. OJIRO, H. NIKI, AND M. USUI, *A new criterion for the H-matrix property*, J. Comput. Appl. Math., 150 (2003), pp. 293–302.
[8] A. OSTROWSKI, *Über die Determinanten mit überwiegender Hauptdiagonale*, Comment. Math. Helv., 10 (1937), pp. 69–96.
[9] R. J. PLEMMONS, *M-matrix characterizations. I. Nonsingular M-matrices*, Linear Algebra Appl., 18 (1977), pp. 175–188.
[10] H. SASSENFELD, *Ein hinreichendes Konvergenzkriterium und eine Fehlerabschätzung für die Iteration in Einzelschritten bei linearen Gleichungen*, Z. Angew. Math. Mech., 31 (1951), pp. 92–94.
[11] R. S. VARGA, *On recurring theorems on diagonal dominance*, Linear Algebra Appl., 13 (1976), pp. 1–9.
[12] R. S. VARGA, *Matrix Iterative Analysis*, expanded ed., Springer, Berlin, 2000.
[13] H. WENDLAND, *Numerical Linear Algebra: An Introduction*, Cambridge University Press, Cambridge, 2017.