HOMOTOPY UNDERSTANDING OF ITERATIVE METHODS

Noah H. Rhee

University of Missouri - Kansas City

Abstract. In this paper we present a homotopy understanding of the iterative methods including the Jacobi, Gauss–Seidel, and Successive Overrelaxation methods. The purpose of this paper is to give a new insight to those iterative methods. Based upon this new understanding we present a condition for the convergence of those iterative methods, which turns out to be equivalent to the traditional one.

1. Introduction. Suppose we want to solve a linear system

$$Ax = b \tag{1.1}$$

by an iterative method. We assume that $A = (a_{ij})$ is a nonsingular m–by–m matrix with complex number entries, and that $b = (b_i) \in \mathbb{C}^m$, that is, b is a vector with m complex number components. We will denote the unique solution of (1.1) by

$$\hat{x} = A^{-1}b. \tag{1.2}$$

Usually such a linear system is solved by a direct method such as finding LU or QR decomposition of A. But direct methods can be impractical if A is large and sparse, because the sought–after factors can be dense. So in such a case we need to study about iterative methods that can solve (1.1).

The Jacobi, Gauss–Seidel, and Successive OverRelaxation (SOR) iterative methods [3,4,5] are typical members of a large family of iterative methods that have the form

$$Mx_{n+1} = Nx_n + b, \quad n = -1, 0, 1, \ldots \tag{1.3}$$

where

$$A = M - N \tag{1.4}$$

is a splitting of A, M is nonsingular and x_{-1} is any vector in $\mathbb{C}^m$. Note that we let n start from -1 (Normally n starts from 0. The reason of this shift of index will become apparent in the next section.)

For the iterative method (1.3) to be practical, it must be easy to solve a linear system with M as the matrix. Henceforth, we will assume that such is the case. We note that for the Jacobi method, M is diagonal, and for the Gauss–Seidel and SOR methods, M is lower triangular.

Whether or not the iterative method (1.3) produces the iterates $\{x_n\}_{n=0}^{\infty}$ that converge to $\hat{x} = A^{-1}b$, which is the solution of (1.1), depends on the eigenvalues of $M^{-1}N$. Of course, the convergence of vectors in $\mathbb{C}^m$ is understood as componentwise convergence. For any $G \in \mathbb{C}^{m\times m}$, the set of all m–by–m matrices whose elements are complex numbers, we define $\rho(G)$ to mean the spectral radius of G, that is, the maximum eigenvalue in magnitude. Then we have the following theorem.

Theorem 1.1. If $\rho(M^{-1}N) < 1$, then the iterates $\{x_n\}_{n=0}^{\infty}$ defined by (1.3), converge to $\hat{x}$, which is the solution of (1.1), for any starting vector $x_{-1} \in \mathbb{C}^m$.

Proof. See [4, p. 59].

In this paper we present a homotopy continuation understanding of the iterative method (1.3). And based upon this new understanding we present a condition for the convergence of the iterative method (1.3), which turns out to be equivalent to the condition $\rho(M^{-1}N) < 1$ which is specified in Theorem 1.1.

2. Homotopy Continuation Approach. Homotopy continuation means that we join a trivial system $Mx = b'$ to the given linear system $Ax = b$ using parameter $t \in [0,1]$. More precisely, we define the homotopy $H\colon \mathbb{C}^m \times [0,1] \to \mathbb{C}^m$ by

$$H(x,t) = A(t)x - b(t) \tag{2.1}$$

where

$$A(t) = M - tN \tag{2.2}$$

$$b(t) = (1-t)Nx_{-1} + b. \tag{2.3}$$

If we assume that $A(t)$ is nonsingular for each $t \in [0,1]$, $H(x,t) = 0$ has a unique solution which we will denote by $\tilde{x}(t)$, for each $t \in [0,1]$. This means that $\tilde{x}$ can be viewed as a complex vector–valued function from $[0,1]$ to $\mathbb{C}^m$. Note that $\tilde{x}(0)$ satisfies the equation $A(0)\tilde{x}(0) - b(0) = 0$. Using the relations (2.2) and (2.3) we obtain

$$M\tilde{x}(0) = Nx_{-1} + b. \tag{2.4}$$

Note that $\tilde{x}(0)$ is immediately available, since a linear system with the matrix M is easy to solve. We also note that $\tilde{x}(1)$ satisfies the equation $A(1)\tilde{x}(1) - b(1) = 0$. Using the relations (2.2), (1.4) and (2.3) we obtain

$$A\tilde{x}(1) = b. \tag{2.5}$$

That is, $\tilde{x}(1) = \hat{x}$, the solution of the given problem.

Now let us study about the complex vector–valued function $\tilde{x}\colon [0,1] \to \mathbb{C}^m$ further. First of all, we need some basic definitions.

Definition 2.1. Let $v: [0,1] \to \mathbb{C}^m$ be a complex vector–valued function of t. We write

$$v(t) = [v_1(t), \dots, v_m(t)]^T$$

Here 'T' stands for 'transpose'. If each component $v_i(t)$ is differentiable for all $t \in [0,1]$, $i = 1, \dots, m$, then we define

$$\frac{d}{dt}v(t) = \left[\frac{d}{dt}v_1(t), \dots, \frac{d}{dt}v_m(t)\right]^T, \quad t \in [0,1].$$

And we say that v is differentiable on $[0,1]$.

If each component $v_i(t)$ is infinitely many times differentiable on $[0,1]$, $i = 1, \dots, m$, then we define for $n = 1, 2, \dots$

$$v^{(n)}(t) = \frac{d^n}{dt^n}v(t) = \left[\frac{d^n}{dt^n}v_1(t), \dots, \frac{d^n}{dt^n}v_m(t)\right]^T, \quad t \in [0,1].$$

And we say that v is infinitely many times differentiable on $[0,1]$. By convention $v^{(0)}(t)$ is understood as $v(t)$.

Definition 2.2. If $v: [0,1] \to \mathbb{C}^m$ is infinitely many times differentiable on $[0,1]$, then for $n = 0, 1, \dots$

$$P_n(v,t) = v(0) + tv^{(1)}(0) + \cdots + \frac{t^n}{n!}v^{(n)}(0), \quad t \in [0,1]$$

is called the nth–degree complex vector–valued Taylor polynomial of $v(t)$ and

$$P_\infty(v,t) = v(0) + tv^{(1)}(0) + \cdots + \frac{t^n}{n!}v^{(n)}(0) + \cdots, \quad t \in [0,1]$$

is called the complex vector–valued Taylor series of $v(t)$.

Definition 2.3. Let $B(t) = (b_{ij}(t)) \in \mathbb{C}^{m \times m}$ and each entry $b_{ij}(t)$ is a differentiable function of t, then we define

$$\frac{d}{dt}B(t) = \begin{pmatrix} \frac{d}{dt}b_{11}(t) & \cdots & \frac{d}{dt}b_{1m}(t) \\ \vdots & \ddots & \vdots \\ \frac{d}{dt}b_{m1}(t) & \cdots & \frac{d}{dt}b_{mm}(t) \end{pmatrix}.$$

If each entry $b_{ij}(t)$ is infinitely many times differentiable, then for $n = 1, 2, \ldots$ we define

$$\frac{d^n}{dt^n}B(t) = \begin{pmatrix} \frac{d^n}{dt^n}b_{11}(t) & \cdots & \frac{d^n}{dt^n}b_{1m}(t) \\ \vdots & \ddots & \vdots \\ \frac{d^n}{dt^n}b_{m1}(t) & \cdots & \frac{d^n}{dt^n}b_{mm}(t) \end{pmatrix}.$$

We also recall the following basic definitions from linear algebra.

Definition 2.4. For any $G = (g_{ij}) \in \mathbb{C}^{m\times m}$, we let $M_{ij}(G)$ be the $(m-1)$–by–$(m-1)$ submatrix of G obtained by deleting row i and column j of G. $M_{ij}(G)$ is called the minor of g_{ij}. And

$$\mathrm{cof}_{ij}(G) = (-1)^{i+j}\det(M_{ij}(G))$$

is called the cofactor of g_{ij}. Here $\det(M_{ij}(G))$ denotes the determinant of the matrix $M_{ij}(G)$.

Definition 2.5. For $G = (g_{ij}) \in \mathbb{C}^{m\times m}$, the adjoint of G, denoted by $\mathrm{Adj}(G) = (g^*_{ij})$, is defined to be the transpose of the matrix obtained from G by replacing each element by its cofactor. That is, $g^*_{ij} = \mathrm{cof}_{ji}(G)$.

Now we state the following basic theorem from linear algebra.

Theorem 2.6. If $G = (g_{ij}) \in \mathbb{C}^{m\times m}$ is nonsingular, then $\det(G) \neq 0$ and

$$G^{-1} = \frac{1}{\det(G)}\mathrm{Adj}(G).$$

Proof. See [2, p. 119].

Now we can state the following lemma.

Lemma 2.7. If the matrix $A(t)$, defined in the relation (2.2), is nonsingular for each $t \in [0,1]$, then the complex vector–valued function $\tilde{x}\colon [0,1] \to \mathbb{C}^m$, defined as the unique solution of $H(x,t) = 0$ (see the relation (2.1) for the definition of $H(x,t)$) for each $t \in [0,1]$, is infinitely many times differentiable on $[0,1]$.

Proof. Let t be any point in $[0,1]$. From the definition of the function $\tilde{x}$ we see that

$$\tilde{x}(t) = A(t)^{-1}b(t). \tag{2.6}$$

By Theorem 2.6 we have

$$A(t)^{-1} = \frac{1}{\det(A(t))}\mathrm{Adj}(A(t)). \tag{2.7}$$

Since each entry of $A(t)$ is linear in the variable t (see (2.2)), from the relation (2.7) and the definitions of adjoint and determinant we see that each entry of $A(t)^{-1}$ is a rational function of t. Furthermore, $\det(A(t))$, which is the common denominator of each entry of $A(t)^{-1}$, is not zero. Hence, each entry of $A(t)^{-1}$ is infinitely many times differentiable.

From the relation (2.3) we see that $b(t)$ is infinitely many times differentiable with respect to the variable t. In fact, we have

$$\frac{d}{dt}b(t) = -Nx_{-1} \quad \text{and} \quad \frac{d^k}{dt^k}b(t) = 0, \quad k = 2, 3, \ldots . \tag{2.8}$$

Hence, from the relation (2.6) we see that $\tilde{x}(t)$ is differentiable. In fact, using the relation (2.8) from the relation (2.6) we obtain

$$\frac{d}{dt}\tilde{x}(t) = \frac{d}{dt}[A(t)^{-1}]b(t) - A(t)^{-1}Nx_{-1}. \tag{2.9}$$

This differentiation procedure can be repeated as many times as we want, since each entry of $A(t)^{-1}$ and $b(t)$ are infinitely many times differentiable. So by applying Leibnitz's rule of differentiation to both sides of the relation (2.6) and using the relation (2.8) we obtain

$$\frac{d^n}{dt^n}\tilde{x}(t) = \frac{d^n}{dt^n}[A(t)^{-1}]b(t) - n\frac{d^{n-1}}{dt^{n-1}}[A(t)^{-1}]Nx_{-1}, \quad n = 1, 2, \ldots . \tag{2.10}$$

Since t was arbitrary, we proved that $\tilde{x}$ is infinitely many times differentiable on $[0, 1]$.

Since the function $\tilde{x}: [0, 1] \to \mathbb{C}^m$ is infinitely many times differentiable, to approximate $\tilde{x}(t)$ we may use the nth–degree complex vector–valued Taylor polynomial $P_n(\tilde{x}, t)$ defined in Definition 2.2 to get for any $t \in [0, 1]$

$$\begin{aligned}\tilde{x}(t) &\approx P_n(\tilde{x}, t)\\ &= \tilde{x}(0) + t\tilde{x}^{(1)}(0) + \cdots + \frac{t^n}{n!}\tilde{x}^{(n)}(0), \quad n = 0, 1, \ldots .\end{aligned}$$

So $\tilde{x}(1) = \hat{x}$, the solution of the given problem (1.1), may be approximated

$$\begin{aligned}\hat{x} = \tilde{x}(1) &\approx P_n(\tilde{x}, 1)\\ &= \tilde{x}(0) + \tilde{x}^{(1)}(0) + \cdots + \frac{1}{n!}\tilde{x}^{(n)}(0), \quad n = 0, 1, \ldots . \end{aligned} \tag{2.11}$$

To manipulate the expression in the relation (2.11) we need the following lemma.

Lemma 2.8. If $A(t)$ is nonsingular for all $t \in [0,1]$, then for any $t \in [0,1]$ we have

$$\tilde{x}^{(1)}(t) = A(t)^{-1}N(\tilde{x}(t) - x_{-1})$$

and

$$\tilde{x}^{(n)}(t) = nA(t)^{-1}N\tilde{x}^{(n-1)}(t), \quad n = 2, 3, \ldots$$

Proof. From the definition of $\tilde{x}(t)$ we see

$$A(t)\tilde{x}(t) = b(t). \tag{2.12}$$

From Lemma 2.7 and the relation (2.2) we know that $\tilde{x}$ and each entry of $A(t)$ are differentiable with respect to the variable t. Hence, differentiating both sides of the relation (2.12) with respect to t we obtain

$$\left[\frac{d}{dt}A(t)\right]\tilde{x}(t) + A(t)\frac{d}{dt}\tilde{x}(t) = \frac{d}{dt}b(t). \tag{2.13}$$

From the relation (2.2) we observe that

$$\frac{d}{dt}A(t) = -N. \tag{2.14}$$

Using the relations (2.8) and (2.14) the relation (2.13) becomes

$$-N\tilde{x}(t) + A(t)\tilde{x}^{(1)}(t) = -Nx_{-1}. \tag{2.15}$$

Hence, we obtain

$$\tilde{x}^{(1)}(t) = A(t)^{-1}N(\tilde{x}(t) - x_{-1}),$$

which proves the first statement of the lemma.

To prove the second statement of the lemma we differentiate the relation (2.15) with respect to the variable t to obtain

$$-N\tilde{x}^{(1)}(t) + \left[\frac{d}{dt}A(t)\right]\tilde{x}^{(1)}(t) + A(t)\tilde{x}^{(2)}(t) = 0.$$

Using the relation (2.14) we obtain

$$\tilde{x}^{(2)}(t) = 2A(t)^{-1}N\tilde{x}^{(1)}(t),$$

which proves the second statement of the lemma with $n = 2$. Suppose the second statement of the lemma is true with $n = k$, that is, we assume

$$\tilde{x}^{(k)}(t) = kA(t)^{-1}N\tilde{x}^{(k-1)}(t).$$

From this equation we obtain

$$A(t)\tilde{x}^{(k)}(t) = kN\tilde{x}^{(k-1)}(t).$$

Differentiating both sides of this equation with respect to the variable t we obtain

$$\left[\frac{d}{dt}A(t)\right]\tilde{x}^{(k)}(t) + A(t)\tilde{x}^{(k+1)}(t) = kN\tilde{x}^{(k)}(t).$$

Using the relation (2.14) this equation becomes

$$\tilde{x}^{(k+1)}(t) = (k+1)A(t)^{-1}N\tilde{x}^{(k)}(t).$$

So the second statement of the lemma is valid with $n = k + 1$. Now the second statement of the lemma follows by mathematical induction.

Repeated use of Lemma 2.8 yields

$$\begin{aligned}
\tilde{x}^{(k)}(t) &= kA(t)^{-1}N\tilde{x}^{(k-1)}(t) \\
&= k(k-1)[A(t)^{-1}N]^2\tilde{x}^{(k-2)}(t) \\
&= \cdots \\
&= k(k-1)\cdots 2[A(t)^{-1}N]^{k-1}\tilde{x}^{(1)}(t) \\
&= k![A(t)^{-1}N]^{k-1}[A(t)^{-1}N(\tilde{x}(t) - x_{-1})] \\
&= k![A(t)^{-1}N]^k(\tilde{x}(t) - x_{-1}).
\end{aligned}$$

So we obtain

$$\frac{1}{k!}\tilde{x}^{(k)}(t) = [A(t)^{-1}N]^k(\tilde{x}(t) - x_{-1}), \quad k = 1, 2, \ldots .$$

Setting $t = 0$ and using the fact that $A(0) = M$, the above equation becomes

$$\frac{1}{k!}\tilde{x}^{(k)}(0) = (M^{-1}N)^k(\tilde{x}(0) - x_{-1}), \quad k = 1, 2, \ldots . \tag{2.16}$$

Using the relation (2.16), the relation (2.11) becomes

$$\begin{aligned}\hat{x} = \tilde{x}(1) &\approx P_n(\tilde{x}, 1)\\ &= \tilde{x}(0) + [M^{-1}N + \cdots + (M^{-1}N)^n](\tilde{x}(0) - x_{-1}) \quad n = 0, 1, \ldots .\end{aligned} \tag{2.17}$$

Now we are ready to state the main theorem in this section.
Theorem 2.9.

$$x_n = P_n(\tilde{x}, 1), \quad n = 0, 1, \ldots .$$

Proof. From the relation (1.3) we have

$$x_0 = M^{-1}(Nx_{-1} + b). \tag{2.18}$$

From the relation (2.4) we note that

$$\tilde{x}(0) = M^{-1}(Nx_{-1} + b). \tag{2.19}$$

Hence, it follows that $x_0 = P_0(\tilde{x}, 1)$. This proves the theorem with $n = 0$. Now suppose that the theorem is true with $n = k$, that is, we assume $x_k = P_k(\tilde{x}, 1)$. We need to prove $x_{k+1} = P_{k+1}(\tilde{x}, 1)$. From the relation (1.3) and using the induction hypothesis and the relations (2.17) and (2.19) we obtain

$$
\begin{aligned}
x_{k+1} &= M^{-1}Nx_k + M^{-1}b \\
&= M^{-1}NP_k(\tilde{x},1) + M^{-1}b \\
&= M^{-1}N\{\tilde{x}(0) + [M^{-1}N + \cdots + (M^{-1}N)^k](\tilde{x}(0) - x_{-1})\} + M^{-1}b \\
&= M^{-1}(N\tilde{x}(0) + b) + [(M^{-1}N)^2 + \cdots + (M^{-1}N)^{k+1}](\tilde{x}(0) - x_{-1}) \\
&= M^{-1}(N\tilde{x}(0) - Nx_{-1} + Nx_{-1} + b) \\
&\quad + [(M^{-1}N)^2 + \cdots + (M^{-1}N)^{k+1}](\tilde{x}(0) - x_{-1}) \\
&= M^{-1}(Nx_{-1} + b) + [(M^{-1}N) + \cdots + (M^{-1}N)^{k+1}](\tilde{x}(0) - x_{-1}) \\
&= \tilde{x}(0) + [(M^{-1}N) + \cdots + (M^{-1}N)^{k+1}](\tilde{x}(0) - x_{-1}) \\
&= P_{k+1}(\tilde{x},1).
\end{aligned}
$$

So the theorem follows by mathematical induction.

Theorem 2.9 tells us that for each $n = 0, 1, \ldots, x_n$, generated by the iterative method (1.3), can be viewed as the nth–degree complex vector–valued Taylor polynomial approximation of $\tilde{x}(1) = \hat{x}$.

3. Convergence Analysis. To find a condition for the convergence of the iterative method (1.3) based upon our homotopy understanding, we need to extend the real homotopy parameter t to a complex parameter z. More precisely, we define the homotopy $H: \mathbb{C}^m \times \mathbb{C} \to \mathbb{C}^m$ such that

$$H(x,z) = A(z)x - b(z) \tag{3.1}$$

where

$$A(z) = M - zN \tag{3.2}$$

and

$$b(z) = (1 - z)Nx_{-1} + b. \tag{3.3}$$

We note that the relations (3.1), (3.2) and (3.3) reduce to the relations (2.1), (2.2) and (2.3), respectively, when $z = t \in [0,1]$. Now we state a lemma.

Lemma 3.1. Suppose $A(z)$ defined in relation (3.2) is nonsingular for all z such that $|z| \leq 1$. Then there exists $\epsilon > 0$ such that $A(z)$ is nonsingular for all z such that $|z| < 1+\epsilon$.

Proof. Let $g(z) = \det(A(z))$. Note that g is a continuous function of z from $\mathbb{C}$ to $\mathbb{C}$, since it is a polynomial of z. Let $h(z) = |g(z)|$. Then clearly h is a continuous function from $\mathbb{C}$ to $\mathbb{R}$. Since the set $F := \{z \ : \ |z| \leq 1\}$ is compact, h attains its extreme values on F. Let e and E be its minimum and maximum value, respectively. Since $A(z)$ is nonsingular on F, $h(z) > 0$ on F. This implies that $e > 0$. Since h is continuous, if we let $U = h^{-1}(e/2, E+1)$, then U is an open set in $\mathbb{C}$ containing the set F. Since F is compact, it follows that there exists $\epsilon > 0$ such that $V := \{z \ : \ |z| < 1+\epsilon\} \subseteq U$. Hence, we have

$$h(V) \subseteq h(U) = h(h^{-1}(e/2, E+1)) \subseteq (e/2, E+1).$$

Since $e/2 > 0$, the above inclusion means that $A(z)$ is nonsingular on V. So the lemma is proved.

If we assume that $A(z)$ is nonsingular for all z such that $|z| \leq 1$, then the homotopy equation $H(x,z) = 0$ (see (3.1)) has a unique solution $\tilde{x}(z)$ for each z such that $|z| \leq 1$. In fact, by Lemma 3.1 such a unique solution $\tilde{x}(z)$ exists for any z such that $|z| < 1+\epsilon$ for some $\epsilon > 0$. So if we let

$$V = \{z \ : \ |z| < 1+\epsilon\}, \tag{3.4}$$

then $\tilde{x}$ can be thought of as a complex vector–valued function from V to $\mathbb{C}^m$.

Now we recall the following basic definition from complex variables [1].

Definition 3.2. A complex–valued function f of the complex variable z is analytic at a point z_0 if its derivative exists not only at z_0 but at each point z in some neighborhood of z_0. It is analytic in an open set U if it is analytic at every point in U.

Now we can state the following theorem.

Theorem 3.3. If $A(z)$ is nonsingular for all z such that $|z| \leq 1$, then each component of the complex vector–valued function $\tilde{x}$ is analytic in V (see (3.4) for the definition of V).

Proof. According to Definition 3.2 all we need to show is that each component of $\tilde{x}$ is differentiable in V; and the proof is entirely similar to that of Lemma 2.7. So we omit the proof.

We recall the following theorem of complex variables.

Theorem 3.4. Let f be analytic inside of a circle C whose center at z_0 and radius r_0. Then at each point z inside C

$$f(z) = f(z_0) + \left[\frac{d}{dz} f(z_0)\right](z - z_0) + \cdots + \frac{1}{n!}\left[\frac{d^n}{dz^n} f(z_0)\right](z - z_0)^n + \cdots,$$

that is, Taylor series converges to $f(z)$ when $|z - z_0| < r_0$.

Proof. See [1, p. 145].

Now we have the following theorem.

Theorem 3.5. Suppose $A(z)$ is nonsingular for all z such that $|z| \leq 1$. Then

$$\tilde{x}(z) = \tilde{x}(0) + z\frac{d}{dz}\tilde{x}(0) + \cdots + \frac{z^n}{n!}\frac{d^n}{dz^n}\tilde{x}(0) + \cdots \tag{3.5}$$

is valid for all z in V.

Proof. The theorem follows from Theorem 3.3 and applying Theorem 3.4 with $z_0 = 0$ and $r_0 = 1 + \epsilon$ to the complex vector–valued function $\tilde{x}$ componentwise. Of course, here the notation

$$\frac{d^k}{dz^k}\tilde{x}^{(k)}(z), \quad k = 1, 2, \ldots$$

is understood as componentwise kth derivative of the complex vector–valued function $\tilde{x}$ with respect to the variable z.

Now we state the main result of this section.

Corollary 3.6. Suppose $A(z)$ is nonsingular for each z such that $|z| \leq 1$. Then the iterates $\{x_n\}_{n=0}^{\infty}$, generated by the iterative method (1.3), converges to $\tilde{x}(1) = \hat{x}$, which is the solution of the given problem (1.1), starting from any vector $x_{-1} \in \mathbb{C}^m$.

Proof. Since $1 \in V$, using Theorem 3.5 we obtain

$$\begin{aligned}
\hat{x} = \tilde{x}(1) &= \tilde{x}(0) + \frac{d}{dz}\tilde{x}(0) + \cdots + \frac{1}{n!}\frac{d^n}{dz^n}\tilde{x}(0) + \cdots \\
&= \tilde{x}(0) + \frac{d}{dt}\tilde{x}(0) + \cdots + \frac{1}{n!}\frac{d^n}{dt^n}\tilde{x}(0) + \cdots \\
&= \tilde{x}(0) + \tilde{x}^{(1)}(0) + \cdots + \frac{1}{n!}\tilde{x}^{(n)}(0) + \cdots \\
&= P_\infty(\tilde{x}; 1) \\
&= \lim_{n\to\infty} P_n(\tilde{x}; 1) \\
&= \lim_{n\to\infty} x_n.
\end{aligned}$$

To get the last equation we used Theorem 2.9. Hence, the corollary is proved.

In Corollary 3.6 we have established a condition, namely, $A(z) = M - zN$ (see (3.2)) is nonsingular for all z such that $|z| \leq 1$, under which the iterates $\{x_n\}_{n=0}^{\infty}$, generated by the iterative method (1.3), converges to $\hat{x} = A^{-1}b$. Now we show that this condition is equivalent to the condition specified in Theorem 1.1, namely, the condition that $\rho(M^{-1}N) < 1$.

Theorem 3.7. The following two statements are equivalent.

(1) $\rho(M^{-1}N) < 1$.

(2) $A(z) = M - zN$ is nonsingular for all z such that $|z| \leq 1$.

Proof. Suppose (1) is true. Suppose that $M - zN$ is singular for some z with $|z| \leq 1$. Then there exists a nonzero vector w such that

$$(M - zN)w = 0. \tag{3.6}$$

Note that $z \neq 0$, since M is nonsingular. So from the relation (3.6) we obtain

$$(M^{-1}N)w = \frac{1}{z}w.$$

This shows that $(1/z, w)$ is an eigenpair of $M^{-1}N$. Since $|1/z| \geq 1$, we have $\rho(M^{-1}N) \geq 1$, which is a contradiction. So we conclude that $A(z) = M - zN$ is nonsingular for all $|z| \leq 1$.

Conversely, suppose (2) is true. Suppose that $\rho(M^{-1}N) \geq 1$. Then there exists an eigenpair (λ, w) of $M^{-1}N$ such that

$$(M^{-1}N)w = \lambda w \quad \text{and} \quad |\lambda| \geq 1.$$

So we have

$$Mw = \frac{1}{\lambda}Nw. \tag{3.7}$$

If we let $z = 1/\lambda$, then $|z| \leq 1$. And from the relation (3.7) we have

$$(M - zN)w = 0 \quad \text{and} \quad w \neq 0.$$

This shows that $M - zN$ is singular for some $|z| \leq 1$, which is a contradiction. Thus, we conclude that $\rho(M^{-1}N) < 1$.

References

1. R. V. Churchill, J. W. Brown, and R. F. Verhey, *Complex Variable and Applications*, (3rd ed.), McGraw–Hill, Inc., New York, 1974.

2. C. G. Cullen, *Matrices and Linear Transformations*, (2nd ed.), Addison–Wesley Publishing Co., Inc. Massachusetts, 1972.

3. L. A. Hageman and D. M. Young, *Applied Iterative Methods*, Academic Press, New York, 1981.

4. R. S. Varga, *Matrix Iterative Analysis*, Prentice–Hall, Englewood Cliffs, N.J., 1962.

5. D. M. Young, *Iterative Solution of Large Linear Systems*, Academic Press, New York, 1971.