

Convergence of iterates of a nonlinear operator arising in statistical mechanics

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1991 Nonlinearity 4 1223

(<http://iopscience.iop.org/0951-7715/4/4/010>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 128.6.62.247

The article was downloaded on 26/09/2010 at 18:14

Please note that [terms and conditions apply](#).

Convergence of iterates of a nonlinear operator arising in statistical mechanics

Roger D Nussbaum†

Mathematics Department, Hill Center, Rutgers University, New Brunswick, NJ 08903, USA

Received 28 August 1990, in final form 12 March 1991

Accepted by J F Toland

Abstract. In studies of an infinite chain of atoms, each atom connected by a ‘spring’ to its nearest neighbours and the whole chain lying in a periodic potential field, several authors have been led to study the nonlinear operator

$$(F_a x)(s) = \min_t [a(s, t) + x(t)]$$

and a finite dimensional version $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ of a F_a given by

$$(T_A(x))_i = \min_j [a_{ij} + x_j].$$

(Here $a(s, t)$ is a given function and A is a given $n \times n$ matrix). The operator T_A also arises in questions from operations research. If $T = T_A$ has a fixed point in \mathbb{R}^n , it is of interest to ask about convergence properties of iterates T^k of T . We shall prove that there is a set $S \subseteq \{j \mid 1 \leq j \leq n\}$ such that $\sum_{j \in S} j \leq n$ and such that if $p = \text{lcm}(S)$, the least common multiple of the integers in S , then $\lim_{k \rightarrow \infty} T^{kp}(x) = \xi(x)$ exists for every $x \in \mathbb{R}^n$. Furthermore, for each $x \in \mathbb{R}^n$ there exists k_x such that $T^{kp}(x) = \xi(x)$ for all $k \geq k_x$. If $T(0) = 0$, we shall provide explicit formulas for p in terms of simple properties of A . Our results are best possible and prove a sharpened version of a conjecture which was made by R B Griffiths in response to a question by this author.

AMS classification scheme numbers: 47H09, 47H10, 58F08, 26A18

PACS number: 0520

Introduction

Before proving our theorems it may be worthwhile to discuss further some of the motivation for studying the operators F_a and T_A defined in the abstract. First, one should note that our methods of proof and theorems are equally valid if ‘min’ in the definition of F_a and T_A is replaced by ‘max’:

$$(\tilde{F}_a x)(s) = \max_t [a(s, t) + x(t)]$$

† Partially supported by NSF DMS 89-03018.

and

$$(\tilde{T}_A(x))_i = \max_j [a_{ij} + x_j].$$

A motivation for studying \tilde{T}_A arises from a problem in ‘machine scheduling’: see [6, chapter 1]. Imagine n machines in a factory. The machines are interconnected and engage in repetitive tasks or ‘cycles’. If machine i completes its r th cycle at time x_i^r , then it begins its $(r + 1)$ th cycle at a time x_i^{r+1} , where

$$x_i^{r+1} = \max_j (a_{ij} + x_j^r).$$

This equation might represent the fact that machine i cannot start its next cycle until receiving output from all or some of the remaining machines. In this interpretation a_{ij} may represent the time needed to take the output of machine j to machine i after cycle r and a_{ii} may represent a necessary cooling period for machine i . Given some initial start-up times x_i^0 , $1 \leq i \leq n$, it is important to understand the long-time behaviour of x_i^r , $1 \leq i \leq n$, as $r \rightarrow \infty$. But this is equivalent to understanding $(\tilde{T}_A)^r(x^0)$, which is exactly the problem addressed in this paper.

We should remark that in practice one may want to allow $a_{ij} = -\infty$ for certain i and j in the definition of \tilde{T}_A , but this case can also be handled by the techniques of this paper.

The connection of maps like T_A and F_a to statistical mechanics arises in the study of so-called Frenkel–Kontorova models. We refer to [4, 11] and especially the survey article [27] for details. Here we must restrict ourselves to a few comments. The map F_a is the object of interest in statistical mechanics, but F_a has been studied by using T_A as a finite-dimensional approximation. In studying F_a or T_A , so-called ‘additive eigenvectors’ play a crucial role. A vector $x \in \mathbb{R}^n$ (respectively $x \in C([0, 1])$) is an additive eigenvector of T_A (respectively F_a) if there exists $\lambda \in \mathbb{R}$ such that

$$(T_A(x))_i = \lambda + x_i \quad 1 \leq i \leq n \quad (\text{respectively } (F_a x)(s) = \lambda + x(s), 0 \leq s \leq 1).$$

Assuming that $a(s, t)$ is continuous, it is known [3] that T_A and F_a have additive eigenvectors. In [8] Floria and Griffiths present a numerical procedure for finding an eigenvector of T_A but present no proof of convergence (see [8, p 566]). A corollary of our approach (see proposition 1.1 and remark 1.1) is a proof of convergence for the Floria–Griffiths method. We should remark that, because F_a is compact for a continuous, our techniques can also be used to prove convergence of the analogous procedure for F_a .

As is discussed in section 3.3 of [27], one is led to study iterates of F_a in considering the ground-state problem for a finite number N of atoms in a periodic potential field $V(u)$, $u \in \mathbb{R}$. If $R_N(u)$ denotes the ‘minimal enthalpy’ of a chain of N atoms at positions u_1, u_2, \dots, u_N , with the constraint that $u_N = u$, then $R_1(u) = V(u)$ and for a suitably defined function $a(s, t)$, $R_N(u) = (F_a^{N-1}(V))(u)$.

Finally, we should remark that the maps T_A provide examples which shed some light on recent nonlinear ergodic theorems [1, 2, 16, 17, 20, 21, 23]. A theorem of Akcoglu and Krengel [1] asserts that if $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is non-expansive with respect to the l_1 norm and f has a fixed point, then for any $x \in \mathbb{R}^n$ there exists some minimal positive integer $p_x = p$ such that $\lim_{k \rightarrow \infty} f^{kp}(x) = \xi$ and $f^p(\xi) = \xi$. It is known that the same theorem is true for any polyhedral norm on \mathbb{R}^n (for example, the sup norm) and that $p_x \leq N$ for all $x \in \mathbb{R}^n$, where N can be chosen to depend only on the polyhedral norm and n . The sup norm is known to play a central role, and it is conjectured in [17] that $N = 2^n$ for the sup norm. The maps T_A are non-expansive

with respect to the sup norm, and theorem 2.2 and remark 2.5 show that in fact, for this particular class of non-expansive maps, much sharper upper bounds on p_x than 2^n can be given.

1. Basic facts about T_A

If A is an $n \times n$ matrix with real entries a_{ij} , we define a map $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ by $(T_A(x))_i =$ the i th component of $T_A(x)$

$$= \min_{1 \leq j \leq n} [a_{ij} + x_j]. \tag{1.1}$$

If x and y are vectors in \mathbb{R}^n , $x \wedge y \equiv z \in \mathbb{R}^n$ is defined by

$$z_i = x_i \wedge y_i = \min(x_i, y_i). \tag{1.2}$$

It is a straightforward exercise to verify that

$$T_A(x \wedge y) = (T_A(x)) \wedge (T_A(y)) \quad \text{for all } x, y \in \mathbb{R}^n. \tag{1.3}$$

We shall write $x \leq y$ for $x, y \in \mathbb{R}^n$ if $y_i - x_i \geq 0$ for $1 \leq i \leq n$, and we shall say that a map $H: D \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ is 'order-preserving on D ' if, for all $x, y \in D$ such that $x \leq y$, it is true that $H(x) \leq H(y)$. Equation (1.3) implies that T_A is order-preserving on \mathbb{R}^n .

We shall always denote by $u \in \mathbb{R}^n$ the vector such that $u_i = 1$ for $1 \leq i \leq n$. In this notation it is easy to show that for all $x \in \mathbb{R}^n$ and all real numbers c

$$T_A(x + cu) = T_A(x) + cu. \tag{1.4}$$

Because T_A is order-preserving and satisfies (1.4), one can prove directly or use a theorem from [5] to conclude that T_A is non-expansive with respect to the sup norm, i.e.,

$$\|T_A(x) - T_A(y)\|_\infty \leq \|x - y\|_\infty \quad \text{for all } x, y \in \mathbb{R}^n \tag{1.5}$$

where

$$\|z\|_\infty = \sup_j |z_j|.$$

If T_A has a fixed point, it follows directly from (1.5) and theorem 1 in [17] that for every $x \in \mathbb{R}^n$ there exists an integer $p_x = p$ such that

$$\lim_{k \rightarrow \infty} T_A^{kp}(x) \text{ exists.}$$

See also [1, 2, 16, 20, 21, 23, 24] for related results.

It is a special case of theorems 1 and 2 in [17] that

$$p_x \leq 2^n \gamma(n)$$

where

$$\gamma(n) \leq n! \left(\frac{1}{\ln(2)} \right)^n.$$

In [20] this estimate on p_x is greatly sharpened by exploiting (1.3), but the methods of [20] cannot, in general, give best possible bounds for the numbers p_x associated with T_A .

Work of Chou and Duffin [3] (see also [7]) indicates exactly when the map T_A has a fixed point. Given an $n \times n$ matrix, define $\lambda = \lambda(A)$ by

$$\lambda = \inf \left\{ \left(\frac{1}{k} \right) \sum_{i=1}^k a_{m_i, m_{i+1}} \mid k \geq 1, 1 \leq m_i \leq n \text{ for } 1 \leq i \leq k \text{ and } m_{k+1} = m_1 \right\} \tag{1.6}$$

Then T_A has a fixed point in \mathbb{R}^n if and only if $\lambda(A) = 0$. In general, there exists $v \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ such that

$$T_A(v) = \lambda u + v \tag{1.7}$$

and the number λ in (1.7) necessarily equals $\lambda(A)$.

These results and corresponding results for the operator F_a mentioned in the abstract can also be obtained (after a change of variables $y_i = e^{x_i}$, $1 \leq i \leq n$) from classical theorems concerning eigenvectors in the interiors of cones for nonlinear operators.

Notice that if v satisfies (1.7), then $v + cu = v'$ also satisfies (1.7), so by choosing c appropriately we can arrange, for any fixed γ , that (1.7) has a solution v with $v_k = \gamma$ for some given k , $1 \leq k \leq n$. It follows that if we define $W_\gamma = \{x \in \mathbb{R}^n \mid x_k = \gamma\}$ and $G : W_\gamma \rightarrow W_\gamma$ by

$$G(x) = T_A(x) - (T_A(x))_k u + \gamma u \tag{1.8}$$

then G has a fixed point v , where v satisfies (1.7) and $v_k = \gamma$.

Define a seminorm $p(x)$ for $x \in \mathbb{R}^n$ by

$$p(x) = \max_{1 \leq i \leq n} x_i - \min_{1 \leq i \leq n} x_i \tag{1.9}$$

and note that the restriction of p to W_0 gives a norm on W_0 and that $p(x + cu) = p(x)$ for all $c \in \mathbb{R}$ and all $x \in \mathbb{R}^n$.

Lemma 1.1. If G and p are defined by equations (1.8) and (1.9) respectively, then for all $x, y \in \mathbb{R}^n$,

$$p(G(x) - G(y)) \leq p(x - y). \tag{1.10}$$

Proof. For x and y in \mathbb{R}^n , define $\alpha = \inf_i(x_i - y_i)$ and $\beta = \sup_i(x_i - y_i)$ and note that

$$y \leq x - \alpha u \leq y + (\beta - \alpha)u. \tag{1.11}$$

We obtain from (1.11) and (1.4) and the fact that T_A is order-preserving that

$$\alpha u \leq T_A(x) - T_A(y) \leq \beta u. \tag{1.12}$$

Equation (1.12) implies that

$$p(T_A(x) - T_A(y)) \leq \beta - \alpha = p(x - y). \tag{1.13}$$

Because $p(z + cu) = p(z)$ for all $z \in \mathbb{R}^n$ and $c \in \mathbb{R}$, we conclude that

$$p(G(x) - G(y)) = p(T_A(x) - T_A(y)) \leq p(x - y). \quad \square$$

Our next proposition gives an iterative procedure for finding a solution v of (1.7). This procedure is similar to one given in [8] (see [8, p 566, equation (1.7)]), but Floria and Griffiths were unable to prove convergence of their iterative scheme. In proposition 1.1 and remark 1.1 below we prove convergence of both our iterative scheme and the Floria–Griffiths scheme.

Proposition 1.1. Let A be a given $n \times n$ matrix, γ a given real number, k an integer with $1 \leq k \leq n$ and $W_\gamma = \{x \in \mathbb{R}^n \mid x_k = \gamma\}$. Let $G : W_\gamma \rightarrow W_\gamma$ be defined by (1.8), and for a fixed h , $0 < h < 1$, define $G_h(x)$ by

$$G_h(x) = (1 - h)x + hG(x). \tag{1.14}$$

If G_h^m denotes the m th iterate of G_h , then for any $x \in W_\gamma$, (a) $G_h^m(x)$ converges to a fixed point v of G and (b) v satisfies (1.7) for $\lambda = \lambda_A$.

Proof. Define a homeomorphism $\varphi : W_0 \rightarrow W_\gamma$ by $\varphi(y) = y + \gamma u$. A calculation shows that

$$(\varphi^{-1}G_h\varphi)(y) = (1 - h)y + h[T_A(y) - (T_A(y))_k u]$$

which corresponds to the map G_h for $\gamma = 0$. Thus in proving convergence of $G_h^m(x)$ to v with $G(v) = v$, it suffices to assume $\gamma = 0$.

Consider $W_0 = W$ with the norm given by $p(x)$ as in (1.9). According to lemma 1.1, G is non-expansive with respect to this norm, and we have remarked that G has a fixed point in W . It follows from a result of Ishikawa [13] that (a) holds. The defining equation for G gives

$$T_A(v) = \alpha u + v$$

for some α , and by the uniqueness of λ in (1.7), $\alpha = \lambda$. □

Remark 1.1. Taking $h = \frac{1}{2}$ and $y \in \mathbb{R}$, we consider iterates of $H = G_{1/2}$, starting with $x \in W_\gamma = \{x \in \mathbb{R}^n \mid x_k = \gamma\}$,

$$H(x) = \frac{1}{2}x + \frac{1}{2}T_A(x) - \frac{1}{2}(T_A(x))_k u + \frac{1}{2}\gamma u.$$

In [8, equation (17)], Floria and Griffiths consider the map $\Gamma : \mathbb{R}^n \rightarrow \mathbb{R}^n$ defined by

$$\Gamma(x) = \frac{1}{2}x + \frac{1}{2}T_A(x) + (\min_i x_i)u - \frac{1}{2}(\min_i (x_i + (T_A(x))_i))u$$

so

$$\min_i (\Gamma x)_i = \min_i x_i.$$

We can easily check that

$$\Gamma(x + cu) = \Gamma(x) + cu \quad \text{and} \quad H(x + cu) = S(x) + \frac{1}{2}cu$$

for all $x \in \mathbb{R}^n$, $c \in \mathbb{R}$. Also we have that

$$\Gamma(x) = H(x) + \theta(x)u,$$

where

$$\theta(x) = \frac{1}{2}(T_A(x))_k + \min_i (x_i) - \frac{1}{2} \min_i (x_i + (T_A(x))_i) - \frac{1}{2}\gamma.$$

Using these equations we can easily prove by induction on m that for all $x \in \mathbb{R}^n$

$$\Gamma^m(x) = H^m(x) + \sum_{j=0}^{m-1} \theta(H^j(x))u = H^m(x) + \alpha_m u. \tag{1.15}$$

If $x \in W_\gamma$, proposition 1.1 implies that $H^m(x) \in W_\gamma$ and $\lim_{m \rightarrow \infty} H^m(x) = v \in W_\gamma$,

where v satisfies (1.7). However, we know

$$\min_i(\Gamma^m(x))_i = \min_i x_i = \min_i(H^m(x))_i + \alpha_m \tag{1.16}$$

and (1.16) implies that α_m converges, so $\Gamma^m(x)$ converges, and the Floria–Griffiths iterative scheme converges for any $x \in \mathbb{R}^n$. This argument also proves that

$$\sum_{j=0}^{\infty} \theta(H^j(x)) = \lim_{m \rightarrow \infty} \alpha_m \text{ exists.}$$

Since $\theta(v) = 0$ for all solutions v of (1.7) and $H^j(x) \rightarrow v$, this fact is plausible but not immediately obvious.

If A is an $n \times n$ matrix and λ and v are as in equation (1.7), define $S(x) = x + v$ (so $S^{-1}(x) = x - v$) for $x \in \mathbb{R}^n$ and define $n \times n$ matrices B and C by $b_{ij} = a_{ij} + v_j - v_i$ and $c_{ij} = b_{ij} - \lambda$. A calculation gives

$$(S^{-1}T_A S)(x) - \lambda u \equiv N(x) = T_B(x) - \lambda u = T_C(x) \tag{1.17}$$

and

$$\min_{1 \leq j \leq n} c_{ij} = 0 \quad \text{for} \quad 1 \leq i \leq n. \tag{1.18}$$

If N is defined by (1.17), one can easily see that

$$N^m(x) = (S^{-1}T_A^m S)(x) - m\lambda u = T_C^m(x)$$

or

$$T_A^m(y) = T_C^m(y - v) + m\lambda u + v. \tag{1.19}$$

It follows that to discuss iterates of T_A it suffices to discuss iterates of T_C , where C satisfies (1.18), so $T_C(0) = 0$. We shall use this observation in the next section.

If σ is a one-one map of $\{j \mid 1 \leq j \leq n\}$ into itself (a permutation), we can associate with σ a linear map $P = P_\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^n$ by $P_\sigma(x) = y$, where $y_j = x_{\sigma^{-1}(j)}$. If we consider elements of \mathbb{R}^n as column vectors and as usual let e_i denote the vector with 1 in row i and zero elsewhere, $P_\sigma(e_i) = e_{\sigma(i)}$ and P_σ , considered as an $n \times n$ matrix (a ‘permutation matrix’), has $e_{\sigma(j)}$ in column j . If A is an $n \times n$ matrix and $B = (b_{ij})$ is defined by $b_{ij} = a_{\sigma(i)\sigma(j)}$, $B = P_\sigma^{-1}AP_\sigma$. Also we have that

$$(T_A P_\sigma x)_i = \min_j (a_{ij} + x_{\sigma^{-1}(j)}) = \min_k (a_{i\sigma(k)} + x_k) \equiv y_i$$

so

$$(P_\sigma^{-1}T_A P_\sigma x)_i = y_{\sigma(i)} = \min_k (a_{\sigma(i)\sigma(k)} + x_k) = (T_{P^{-1}AP}x)_i.$$

Thus, if P is a permutation matrix, we have

$$P^{-1}T_A P = T_{P^{-1}AP}, \tag{1.20}$$

which we shall need later.

If A and B are $n \times n$ matrices, it is a simple calculation to see (or use results in [7]) that

$$T_B T_A = T_{B * A} \tag{1.21}$$

where

$$B * A = C \quad \text{and} \quad c_{ik} = \min_j (b_{ij} + a_{jk}) \tag{1.22}$$

Because of equations (1.19) and (1.18) we will be working with non-negative matrices A , and it will be useful to associate with each such matrix another non-negative matrix B .

Definition 1.1. Let A be an $n \times n$ matrix with $a_{ij} \geq 0$ for all i and j . Say that an $n \times n$ matrix B is 'associated with A ' or that ' A and B are associated' if $a_{ij} > 0$ implies that $b_{ij} = 0$ and $a_{ij} = 0$ implies that $b_{ij} > 0$.

Equivalently, non-negative $n \times n$ matrices A and B are associated iff $a_{ij}b_{ij} = 0$ and $a_{ij} + b_{ij} > 0$ for all i and j . By using this second characterization, it is easy to see that if A and B are associated and P is a permutation matrix then $P^{-1}AP$ and $P^{-1}BP$ are associated.

Lemma 1.2. Suppose that A and B are $n \times n$ non-negative matrices and that C is a matrix associated with A and D is associated with B . Then DC is associated with $B * A$ (see (1.22)). If $(T_A)^m = T_F$ and F has entries f_{ij} , then $f_{ij} = 0$ if and only if $c_{ij}^{(m)} > 0$, where $c_{ij}^{(m)}$ is the entry in row i , column j of C^m .

Proof. If $E = B * A$, we have

$$e_{ik} = \min_j (b_{ij} + a_{jk})$$

so $e_{ik} = 0$ if and only if there exists j such that $b_{ij} = 0$ and $a_{jk} = 0$. However this happens if and only if there exists j such that $d_{ij} > 0$ and $c_{jk} > 0$, and (since all entries of D and C are non-negative), this is equivalent to

$$\sum_p d_{ip}c_{pk} > 0.$$

This shows $e_{ik} = 0$ if and only if the i, k entry of DC is positive, so DC and $B * A$ are associated.

By repeatedly using (1.21) and the first part of this lemma we see that F is associated with C^m , which proves the second part of the lemma. □

2. Convergence of $T_A^{kp}(x)$ as $k \rightarrow \infty$: the optimal p

Before discussing T_A we shall have to consider a closely related family of mappings. We begin with a simple, general lemma, which actually holds for $C(S)$, S a compact, Hausdorff space, not just for \mathbb{R}^m .

Lemma 2.1. Suppose that $f: \mathbb{R}^m \rightarrow \mathbb{R}^m$ and $g: \mathbb{R}^m \rightarrow \mathbb{R}^m$ are non-expansive with respect to the sup norm. Then $h = f \wedge g$ (so $h(x) = f(x) \wedge g(x)$) is non-expansive with respect to the sup norm. If f and g are order-preserving, h is order-preserving.

Proof. We leave the second part of the lemma to the reader. To prove the first part, suppose that $x, y \in \mathbb{R}^m$ and $\|x - y\| = d$. For a fixed i , define $f_i(x) = a_i$, $f_i(y) = b_i$, $g_i(x) = \alpha_i$ and $g_i(y) = \beta_i$. We know that

$$|a_i - b_i| \leq d \quad \text{and} \quad |\alpha_i - \beta_i| \leq d$$

and we have to prove that $|\alpha_i \wedge a_i - \beta_i \wedge b_i| \leq d$.

By symmetry in the roles of f and g , we can assume that $a_i \wedge \alpha_i = a_i$. If $b_i \wedge \beta_i = b_i$, we have

$$|\alpha_i \wedge a_i - \beta_i \wedge b_i| = |a_i - b_i| \leq d$$

and we are done. Thus we can assume $\beta_i \wedge b_i = \beta_i$. Then we have

$$\beta_i \leq b_i \leq a_i + d \quad \text{and} \quad \beta_i \geq \alpha_i - d \geq a_i - d,$$

so

$$|\beta_i - a_i| = |\alpha_i \wedge a_i - \beta_i \wedge b_i| \leq d. \quad \square$$

We introduce a slight generalization of the maps T_A . Suppose that $c \in \mathbb{R}^n$ and A is an $n \times n$ real matrix and $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is defined by

$$(f(x))_i = \min_{1 \leq j \leq n} (c_i, a_{ij} + x_j) = c_i \wedge \min_j (a_{ij} + x_j). \tag{2.1}$$

Definition 2.1. We will say that f satisfies H2.1 if there exist $c \in \mathbb{R}^n$ and an $n \times n$ real matrix A such that $a_{ij} \geq 0$ for all i, j and such that $f(x)$ is given by (2.1).

Definition 2.2. We will say that f satisfies H2.2 if f satisfies H2.1 and $a_{ij} > 0$ for $i \leq j$.

Lemma 2.2. If f and g satisfy H2.1, $h(x) = g(f(x))$ satisfies H2.1, and if f and g satisfy H2.2, h satisfies H2.2. If f satisfies H2.1, then f is non-expansive with respect to the sup norm and f has a fixed point. Furthermore, for each $x \in \mathbb{R}^n$, there exists an integer $p = p_x$ such that $\lim_{k \rightarrow \infty} f^{kp}(x)$ exists.

Proof. If f and g satisfy H2.1 we have

$$y_j = (f(x))_j = \min_k (c_j, a_{jk} + x_k)$$

where $a_{ij} \geq 0$ for all i and j and

$$(g(y))_i = \min_j (d_i, b_{ij} + y_j).$$

It follows that

$$\begin{aligned} (g(f(x)))_i &= (h(x))_i = \min_j (d_i, b_{ij} + \min_k (c_j, a_{jk} + x_k)) \\ &= \min_j \min_k (d_i, c_j + b_{ij}, a_{jk} + b_{ij} + x_k) \\ &= \min_k (d_i, \min_j (c_j + b_{ij}), x_k + \min_j (b_{ij} + a_{jk})). \end{aligned}$$

If we define $\beta_i = \min_j (c_j + b_{ij})$, $\gamma_i = d_i \wedge \beta_i$ and $\delta_{ik} = \min_j (b_{ij} + a_{jk})$, the i, k entry of $B * A$, then

$$(g(f(x)))_i = \min_k (\gamma_i, \delta_{ik} + x_k)$$

and the matrix associated with h is $B * A$. The formula for δ_{ik} shows $\delta_{ik} \geq 0$, because A and B have only non-negative entries. Thus h satisfies H2.1.

If f and g satisfy H2.2, we have to show that $\delta_{ik} > 0$ for $k \geq i$, or, equivalently that $b_{ij} + a_{jk} > 0$ for $1 \leq j \leq n$ and $k \geq i$. If $j \leq k$ we have

$$b_{ij} + a_{jk} \geq a_{jk} > 0$$

and if $j > k \geq i$, we have

$$b_{ij} + a_{jk} \geq b_{ij} > 0.$$

This proves h satisfies H2.2. (This result also follows immediately from lemma 1.2).

If f is as defined before, let $\varphi(x) = c \in \mathbb{R}^n$, where c_i is as in the definition of f . Then we have $f = T_A \wedge \varphi$. We already know that T_A is non-expansive with respect to the sup norm and φ obviously is, so lemma 2.1 implies that f is non-expansive. Similarly, lemma 2.1 implies that f is order-preserving.

To prove that f has a fixed point, suppose, that f and c are as above and define

$$\kappa = \min_i c_i.$$

By definition we have that $f(x) \leq c$ for all $x \in \mathbb{R}^n$. On the other hand, if $\kappa u \leq x \leq c$, we have

$$(f(x))_i = \min_j (c_i, a_{ij} + x_j) \geq \min_j (\kappa, a_{ij} + \kappa) \geq \kappa$$

where we have used the fact $a_{ij} \geq 0$ for all i, j . If $D = \{x \mid \kappa u \leq x \leq c\}$, D is a compact, convex set and $f(D) \subset D$, so f has a fixed point in D .

Given that f is non-expansive and f has a fixed point, the final assertion of the lemma follows immediately from theorem 1 in [17]. □

For our later discussion of the map T_A , we shall need a theorem concerning maps f which satisfy H2.2 (see definition 2.2). Since no more work is involved, we shall actually prove a result about a more general class of maps.

We are indebted to Xianwen Xie for the following proof, which simplifies an earlier argument of this author.

Theorem 2.1. Let Q be a compact subset of \mathbb{R} and $a: Q \times Q \rightarrow \mathbb{R}$ a continuous map such that $a(s, t) \geq 0$ for all $s, t \in Q$ and $a(s, t) > 0$ for all $s, t \in Q$ such that $s \leq t$. Let $c: Q \rightarrow \mathbb{R}$ be a continuous map. If $X = C(Q)$, the Banach space of continuous real-valued functions in the sup norm, define a map $F: X \rightarrow X$ by

$$(Fx)(s) = c(s) \wedge \min_{t \in Q} [a(s, t) + x(t)].$$

Then F has a unique fixed point ξ in X and for every $x \in X$,

$$\lim_{k \rightarrow \infty} F^k(x) = \xi.$$

In particular (the case $Q = \{j \in \mathbb{Z} \mid 1 \leq j \leq m\}$), if $f: \mathbb{R}^m \rightarrow \mathbb{R}^m$ satisfies H2.2, then f has a unique fixed point $\xi \in \mathbb{R}^m$ and

$$\lim_{k \rightarrow \infty} f^k(x) = \xi \quad \text{for all } x \in \mathbb{R}^m.$$

Proof. Essentially the same argument used in [3] shows that F takes bounded sets to precompact sets, so F is a compact map. If maps $F_j: X \rightarrow X$, $j = 1, 2$, are defined by

$$(F_1(x))(s) = c(s) \quad \text{and} \quad (F_2(x))(s) = \min_{t \in Q} [a(s, t) + x(t)]$$

F_1 is obviously order-preserving and non-expansive in the sup norm, and the same arguments used in section 1 show that F_2 is order-preserving and non-expansive in the sup norm. The proof of lemma 2.1 now shows that $F = F_1 \wedge F_2$ is non-expansive in the sup norm and order preserving.

Let $\kappa = \min_{s \in Q} c(s)$. If $F(x) = x$, we claim that

$$\kappa \leq x(s) \leq c(s) \quad \text{for } s \in Q. \tag{2.2}$$

Since $(Fx)(s) \leq c(s)$ for all $x \in X$, it suffices to prove $x(s) \geq \kappa$. Suppose not, so that $\kappa' = \min_{s \in Q} x(s) < \kappa$ and let $s_0 = \min\{s \in Q \mid x(s) = \kappa'\}$. Then we have

$$\begin{aligned} x(s_0) &= \kappa' = c(s_0) \wedge \min_{t \in Q} [a(s_0, t) + x(t)] \\ &= \min_{t \in Q} [a(s_0, t) + x(t)]. \end{aligned} \tag{2.3}$$

However $x(t) > \kappa'$ if $t < s_0$ and $a(s_0, t) > 0$ if $t \geq s_0$, so the right-hand side of (2.3) is strictly greater than κ' , a contradiction.

If $J = \{x \in X \mid \kappa \leq x(s) \leq c(s) \text{ for all } s\}$, it is clear that $F(J) \subset J$, so $F(\kappa u) \geq \kappa u$ (where u denotes the function $u(s) = 1$ for $s \in Q$) and $F(c) \leq c$. Because F is compact and order-preserving, it follows that $x_n = F^n(\kappa u)$ is an increasing sequence of functions in X and $\{x_n \mid n \geq 0\}$ is precompact. Thus

$$\lim_{n \rightarrow \infty} F^n(\kappa u) = \underline{x} \quad \text{exists and } F(\underline{x}) = \underline{x}.$$

A similar argument shows that

$$\lim_{n \rightarrow \infty} F^n(c) = \bar{x} \geq \underline{x} \quad \text{exists and } F(\bar{x}) = \bar{x}.$$

Suppose that we can prove that $\bar{x} = \underline{x} = \xi$. Then if x is any fixed point of F , $\kappa u \leq x \leq c$ and

$$F^n(\kappa u) \leq F^n(x) = x \leq F^n(c)$$

which, after taking limits, implies that $x = \xi$.

Thus, to prove uniqueness of the fixed point of F , it suffices to prove $\underline{x} = \bar{x}$. We suppose not, so $\alpha = \|\bar{x} - \underline{x}\| > 0$ and obtain a contradiction. If $s_j \in Q$ is such that

$$\bar{x}(s_j) - \underline{x}(s_j) = \alpha$$

we know that $\underline{x}(s_j) < c(s_j)$. For if $\underline{x}(s_j) = c(s_j)$, we have $c(s_j) = \underline{x}(s_j) < \bar{x}(s_j)$ and $\bar{x}(s_j) = c(s_j) \wedge \min_{t \in Q} [a(s_j, t) + \bar{x}(t)] \leq c(s_j)$, a contradiction. Thus we obtain that

$$\underline{x}(s_j) = \min_{t \in Q} [a(s_j, t) + \underline{x}(t)]$$

Select $s_{j+1} \in Q$ such that

$$\underline{x}(s_j) = a(s_j, s_{j+1}) + \underline{x}(s_{j+1}) \tag{2.4}$$

Of course we also have

$$\bar{x}(s_j) \leq a(s_j, s_{j+1}) + \bar{x}(s_{j+1}). \tag{2.5}$$

If strict inequality held in (2.5) we would find that

$$\alpha = \bar{x}(s_j) - \underline{x}(s_j) = \|\bar{x} - \underline{x}\| < \bar{x}(s_{j+1}) - \underline{x}(s_{j+1}) \leq \alpha$$

a contradiction. Thus equality holds in (2.5) and

$$\alpha = \bar{x}(s_{j+1}) - \underline{x}(s_{j+1}) = \bar{x}(s_j) - \underline{x}(s_j). \tag{2.6}$$

By virtue of (2.6) we see that $s_k \in Q$ can be defined for all $k \geq 0$ and

$$\alpha = \bar{x}(s_k) - \underline{x}(s_k) \quad \text{and} \quad \underline{x}(s_k) = a(s_k, s_{k+1}) + \underline{x}(s_{k+1}).$$

By applying (2.4) repeatedly we see that

$$\underline{x}(s_0) = \sum_{j=0}^n a(s_j, s_{j+1}) + \underline{x}(s_{n+1})$$

and since $\underline{x}(s_{n+1})$ is bounded and $a(s_j, s_{j+1}) \geq 0$ we conclude that

$$\sum_{j=0}^{\infty} a(s_j, s_{j+1}) < \infty. \tag{2.7}$$

Select a positive number δ such that $a(s, t) \geq \delta$ if $s, t \in Q$ and $s \leq t$. It follows that $a(s_j, s_{j+1}) \geq \delta$ if $s_{j+1} \geq s_j$, so (2.7) implies that there exists an integer N such that $s_{j+1} < s_j$ for all $j \geq N$. It follows that $\lim_{j \rightarrow \infty} s_j = s$ exists. Using (2.4) and the continuity of \underline{x} and a we find

$$\underline{x}(s) = a(s, s) + \underline{x}(s)$$

which contradicts the fact that $a(s, s) > 0$.

It remains to prove that $\lim_{j \rightarrow \infty} F^j(x) = \xi$ for all $x \in X$. By replacing x by Fx , we can assume that $x \leq c$. If $x \geq \kappa u$ (where $\kappa = \min, c(s)$), we obtain that

$$F^n(\kappa u) \leq F^n(x) \leq F^n(c)$$

which implies the desired result. Thus assume that

$$\min_{s \in Q} x(s) = \kappa' < \kappa.$$

However, one can easily see that $F(\kappa' u) \geq \kappa' u$, so $F^j(\kappa' u)$ is an increasing sequence of continuous functions. Since $F^j(\kappa' u)$ is bounded above, $\{F^j(\kappa' u) : j \geq 0\}$ is precompact $\lim_{j \rightarrow \infty} F^j(\kappa' u)$ exists and equals the unique fixed point ξ of F . We conclude that

$$F^j(\kappa' u) \leq F^j(x) \leq F^j(c) \quad \text{for } j \geq 1$$

so $F^j(x)$ also converges to ξ . □

An examination of the proof of theorem 2.1 shows that the fact that Q is a subset of \mathbb{R} plays a limited role. It suffices that Q be a compact metric space, that Q be totally ordered by an ordering $s \leq t$ and that the ordering be continuous, i.e., $s_n \rightarrow s, t_n \rightarrow t$ and $s_n \leq t_n$ for all n implies $s \leq t$.

Theorem 2.1 will play an important role in proving our main result, theorem 2.2.

To proceed further we need to recall some basic facts about non-negative irreducible matrices. In order to keep this paper as self-contained as possible, we state the basic definitions and sketch some proofs.

If $A = (a_{ij})$ is an $n \times n$ non-negative matrix (so $a_{ij} \geq 0$ for all i and j), A is called 'irreducible' if, for each pair of integers i and j such that $i \neq j$ and $1 \leq i, j \leq n$, there is a positive integer $m = m(i, j)$, such that $a_{ij}^{(m)}$, the entry in row i and column j of A^m , is positive. Notice that in this definition, the 1×1 zero matrix (0) is irreducible.

Equivalently, an $n \times n$ non-negative matrix A is irreducible if $n = 1$ or if $n > 1$ and there does *not* exist a permutation matrix P such that

$$P^{-1}AP = \begin{bmatrix} B & 0 \\ D & C \end{bmatrix}$$

where B and D are square matrices (see [10, 15, 22]).

If B is an $n \times n$ non-negative matrix, then B can be written in ‘Perron–Frobenius normal form’, i.e., there is a permutation matrix P such that

$$P^{-1}BP = \begin{bmatrix} B_{11} & & & 0 \\ B_{21} & B_{22} & & \\ & & \ddots & \\ B_s & & & B_{ss} \end{bmatrix} \tag{2.8}$$

where B_{ii} , $1 \leq i \leq s$, is a square, irreducible matrix.

If B is an $m \times m$ non-negative, irreducible matrix and B is not the zero matrix, there is a positive integer $d = d(B) \leq m$ and an integer k_0 such that all elements on the main diagonal of B^{kd} are positive for all $k \geq k_0$. This fact is a special case of a more general classical theorem: see, for example, theorem 4.8 on p 153 in [10]. Because this result is fairly easy and because it plays a central role in our proof of theorem 2.2, we sketch a proof.

If S is a finite set of positive integers, $\text{lcm}(S)$ and $\text{gcd}(S)$ denote the least common multiple of the elements of S and the greatest common divisor of the elements of S . If S is a (possibly infinite) set of positive integers we define

$$\text{gcd}(S) = \lim_{n \rightarrow \infty} \text{gcd}(S_n) \quad \text{where } S_n = \{j \in S \mid j \leq n\}.$$

If S is a set of positive integers which is closed under addition ($\alpha, \beta \in S$ implies $\alpha + \beta \in S$) and $d = \text{gcd}(S)$, it is an easy exercise to prove (see the appendix in [10]) that there exists an integer k_0 such that $kd \in S$ for all $k \geq k_0$.

If B is an $m \times m$, non-negative, irreducible matrix, $B \neq (0)$, and $1 \leq j \leq m$, let $S_j = \{r \geq 1 \mid b_{jj}^{(r)} > 0\}$ ($b_{jj}^{(r)}$ is the j, j entry of B^r) and $d_j = \text{gcd}(S_j)$. It is not hard to prove that the least element of S_j is less than or equal to m , so $d_j \leq m$. It is also true that S_j is closed under addition: if $b_{jj}^{(r)} > 0$ and $b_{jj}^{(s)} > 0$, then $b_{jj}^{(r+s)} \geq b_{jj}^{(r)}b_{jj}^{(s)} > 0$. It follows that there exists an integer k_0 such that $kd_j \in S_j$ for all integers $k \geq k_0$ and for $1 \leq j \leq m$. We claim that $d_i = d_j$ for $i \neq j$. For notational convenience put $d = d_i$ and $d' = d_j$. We know (by irreducibility of B) that there exist positive integers s and t such that $b_{ij}^{(s)} > 0$ and $b_{ji}^{(t)} > 0$. Thus for all $k \geq k_0$ we have

$$b_{jj}^{(kd+s+t)} \geq b_{ji}^{(t)}b_{ii}^{(kd)}b_{ij}^{(s)} > 0.$$

It follows that, for $k \geq k_0$, $(k + 1)d + s + t$ is an integer multiple of d' and $kd + s + t$ is an integer multiple of d' , so d is an integer multiple of d' . This implies $d \geq d'$, and by a symmetric argument we get $d' \geq d$ and, finally, $d = d'$.

Definition 2.3. If B is an $m \times m$ non-negative irreducible matrix, define $d(B) = 1$ if $m = 1$ and, for $m > 1$ and $1 \leq i \leq m$, $d(B) = \text{gcd}(\{r \mid b_{ii}^{(r)} > 0\})$.

Our remarks above show that $d(B) = d \leq m$, that $d(B)$ is independent of i in definition 2.3 and that, if B is not the zero matrix, there is an integer k_0 such that all the entries on the main diagonal of B^{kd} are positive for $k \geq k_0$.

for all positive integers k . If $(T_A^p(\xi))_i < \xi_i$ for some $i \notin J$, then by using (2.13) repeatedly we find that

$$(T_A^{kp}(\xi))_i < \xi_i \quad \text{for all integers } k \geq 1. \tag{2.14}$$

Equation (2.14) contradicts the fact that $T_A^{kp}(\xi) = \xi$ for $k = k_0q$, and we conclude that

$$(T_A^p(\xi))_i = \xi_i \quad \text{for all } i \notin J. \tag{2.15}$$

For notational convenience, define $T_A^p = T_C$ and $\xi^{(k)} = T_A^{kp}(\xi) = T_C^k(\xi)$. We know that B^p is associated with C , so $b_{ij}^{(p)} = 0$ for $i \in J$ and $j \geq i$ and $c_{ij} > 0$ for $i \in J$ and $j \geq i$. We know that $\xi_i^{(k)} = \xi_i$ for all $k \geq 0$ and all $i \notin J$. Thus we can write, for $i \in J$,

$$\xi_i^{(k+1)} = (T_C(\xi^{(k)}))_i = \min_{j \notin J} (c_{ij} + \xi_j) \wedge \min_{j \in J} (c_{ij} + \xi_j^{(k)}).$$

If we define α_i by

$$\begin{aligned} \alpha_i &= \min_{j \notin J} (c_{ij} + \xi_j) \\ \xi_i^{(k+1)} &= \alpha_i \wedge \min_{j \in J} (c_{ij} + \xi_j^{(k)}) \quad \text{for } i \in J. \end{aligned} \tag{2.16}$$

Let μ denote the number of elements in J (if $\mu = 0$ we are done) and identify \mathbb{R}^μ with the $z \in \mathbb{R}^n$ such that $z_i = 0$ for $i \notin J$. Define a map $G: \mathbb{R}^\mu \rightarrow \mathbb{R}^\mu$ by $(G(z))_i = 0$ for $i \notin J$ and

$$(G(z))_i = \alpha_i \wedge \min_{j \in J} (c_{ij} + z_j) \quad \text{for } i \in J. \tag{2.17}$$

Because $c_{ij} \geq 0$ for all $i, j \in J$ and $c_{ij} > 0$ for $i \in J$ and all $j \geq i$, G satisfies the conditions of theorem 2.1. Thus, for any $z \in \mathbb{R}^\mu$, $\lim_{k \rightarrow \infty} G^k(z) = \zeta$ exists, where $\zeta \in \mathbb{R}^\mu$ is the unique fixed point of G . On the other hand, if z is chosen so $z_j = \xi_j$ for all $j \in J$, (2.16) and (2.17) imply that

$$\xi_i^{(k)} = (G^k(z))_i \quad \text{for all } i \in J. \tag{2.18}$$

If $k = mq$ we conclude from (2.18) that

$$\xi_i^{(mq)} = \xi_i = \lim_{m \rightarrow \infty} (G^{mq}(z))_i = \zeta_i. \tag{2.19}$$

Therefore we have $z = \zeta$ and $G^k(z) = G^k(\zeta)$ for all $k \geq 1$, so (2.18) implies that $T_A^p(\xi) = \xi$. □

Remark 2.1. If the matrix B in theorem 2.2 is irreducible and $d = d(B)$ (see definition 2.3), so $d \leq n$, then $\lim_{k \rightarrow \infty} T_A^{kd}(x)$ exists for every $x \in \mathbb{R}^n$.

Remark 2.2. Suppose that A is an $n \times n$ non-negative matrix and that there exists a permutation σ of $M = \{j \mid 1 \leq j \leq n\}$ such that $a_{ij} = 0$ if and only if $j = \sigma(i)$. If $\delta = \inf\{a_{ij} \mid a_{ij} > 0\}$ and $x \in \mathbb{R}^n$ is such that $0 < x_i < \delta$ for $1 \leq i \leq n$, then $T_A(x) = y$ where $y_i = x_{\sigma(i)}$. Thus $T_A^k(x) = P^k(x)$, where P is the permutation matrix corresponding to σ^{-1} . By choosing the permutation σ appropriately, the number p in theorem 2.2 can be taken equal to $\text{lcm}(S)$, for any set $S \subset \{j \mid 1 \leq j \leq n\}$ such that $\sum_{j \in S} j \leq n$. This observation has also been made by Griffiths [12].

The case of a general matrix A follows trivially from theorem 2.2.

Theorem 2.3. Let A be an $n \times n$ matrix, let λ be given by (1.6) and v be a solution of (1.7). Let C be the $n \times n$ matrix given by $c_{ij} = a_{ij} + v_j - v_i - \lambda$ and $p = \theta(D)$ (as in (2.9)), where D is a non-negative matrix associated with C . Then for every $x \in \mathbb{R}^n$,

$$\lim_{k \rightarrow \infty} (T_A^{kp}(x) - kp\lambda u)$$

exists

Proof. This follows immediately from theorem 2.2 and (1.19). □

Remark 2.3. If B is an $n \times n$ non-negative matrix, the computation of $\theta(B)$ in (2.9) seems to require putting B into Perron–Frobenius normal form, but this is not the case. For each i , $1 \leq i \leq n$, let $S_i = \{r \geq 1 \mid b_{ii}^{(r)} > 0\}$ and let $J = \{i \mid S_i \text{ is empty}\}$. If $i \notin J$, let $d_i = \text{gcd}(S_i)$. We leave to the reader the argument that $\theta(B) = \text{lcm}(\{d_i \mid i \notin J\})$. Because S_i is closed under addition, there exists k_0 such that for $\theta = \theta(B)$ and for all $k \geq k_0$,

$$b_{ii}^{(k\theta)} > 0 \quad \text{for all } i \notin J.$$

From this characterization of $\theta(B)$ we easily see that if B_1, B_2 and $B_2 - B_1$ are non-negative $n \times n$ matrices, then $\theta(B_2)$ divides $\theta(B_1)$.

Our final result shows that the convergence of iterates $T_A^{kp}(x)$ in theorem 2.3 actually stabilizes after only finitely many steps for any given x .

Theorem 2.4. Let A be an $n \times n$ matrix such that $\lambda(A) = 0$ (see (1.6)), so T_A has a fixed point. Let p be defined as in theorem 2.3. Then for each $\bar{x} \in \mathbb{R}^n$, there exists an integer $k(\bar{x})$ such that $T_A^{kp}(\bar{x}) = \xi$ for all $k \geq k(\bar{x})$ and $T_A^p(\xi) = \xi$.

Proof. Let C be a matrix such that $T_A^p = T_C$. We know that for any $\bar{x} \in \mathbb{R}^n$, $\lim_{k \rightarrow \infty} T_C^k(\bar{x}) = \xi$, where $T_C(\xi) = \xi$. If we define $S(x) = x + \xi$, we have that

$$S^{-1}T_C S = T_D \quad \text{where } d_{ij} = c_{ij} + \xi_j - \xi_i$$

and $T_D(0) = 0$. We also know that

$$\lim_{k \rightarrow \infty} T_D^k(\bar{x} - \xi) = \lim_{k \rightarrow \infty} S^{-1}T_C^k(\bar{x}) = 0.$$

If we define $\bar{y} = \bar{x} - \xi$, it suffices to prove that there exists an integer $k(\bar{x})$ such that

$$T_D^k(\bar{y}) = 0 \quad \text{for all } k \geq k(\bar{x}).$$

Define $\delta > 0$ by

$$2\delta = \min\{d_{ij} \mid d_{ij} > 0\}.$$

Let $U = \{y \in \mathbb{R}^n \mid |y_i| < \delta \text{ for } 1 \leq i \leq n\}$. Because $T_D(cu) = cu$ for all $c \in \mathbb{R}$ and because T_D is order-preserving, we have $T_D(u) \subset U$. If we define $S_i = \{j \mid d_{ij} = 0\}$, one can see that

$$(T_D(y))_i = \min_{j \in S_i} (y_j) \equiv (H(y))_i \quad \text{for } y \in U$$

because $d_{ij} + y_j > \delta$ for $y \in U$ and $j \notin S_i$.

Select k_0 so that $z = T_D^{k_0}(\bar{y}) \in U$. The above remarks show that

$$(T_D^m(z))_i = (H^m(z))_i = \min_{j \in S(i,m)} (z_j) \tag{2.20}$$

where $S(i, m)$ is a non-empty subset of $\{j \mid 1 \leq j \leq n\}$ depending on i and m . There are only finitely many vectors of the form given by the right-hand side of (2.20), so there must be integers $1 \leq k < m$ such that

$$T_D^m(z) = T_D^k(z).$$

It follows that for every integer $\alpha > 1$,

$$T_D^{\alpha(m-k)+k}(z) = T_D^k(z). \tag{2.21}$$

Since the left-hand side of (2.21) approaches 0 as $\alpha \rightarrow \infty$,

$$T_D^k(z) = 0. \tag{□}$$

Remark 2.4. It may be helpful to discuss the form which theorem 2.4 takes for the example in remark 2.2. If notation is as in remark 2.2 and p denotes the order of the permutation σ , our results imply that $T_A^p = T_C$, where C is an $n \times n$ non-negative matrix such that $c_{ij} > 0$ for $i \neq j$ and $c_{ii} = 0$. We shall only need the fact that $c_{ii} = 0$. Theorem 2.4 implies that for any $x \in \mathbb{R}^n$, there exists k_x such that $T_C^k(x) = \xi$ for all $k \geq k_x$ and $T_C(\xi) = \xi$. However, more can be said. If $x \in \mathbb{R}^n$, let $x^j = T_C^j(x)$. Let $K_1 = \{i \mid x_i = \min_s x_s\}$ and, for $i \in K_1$, let $x_i = \mu_1$. If x_i^j denotes the i th component of x^j , one can show that $x_i^j = x_i$ for $j \geq 1$ and $i \in K_1$. Define $K_2 = \{i \notin K_1 : x_i^1 = \min_{s \notin K_1} x_s^1\}$ and for $i \in K_2$, let $x_i^1 = \mu_2$. Then one can see that $\mu_2 \geq \mu_1$ and $x_i^j = x_i^1$ for $j \geq 2$ and $i \in K_2$. In general, proceed inductively: if K_i and μ_i have been defined for $1 \leq i \leq m$, write $\Gamma_m = \bigcup_{i=1}^m K_i$ and, if $\Gamma_m \neq \{i \mid 1 \leq i \leq n\}$, define $K_{m+1} = \{i \notin \Gamma_m : x_i^m = \min_{s \notin \Gamma_m} x_s^m\}$ and $\mu_{m+1} = x_i^m$ for $i \in K_{m+1}$. One can prove that $x_i^j = x_i^m$ for $j \geq m + 1$ and $i \in K_{m+1}$ and that $\mu_{m+1} \geq \mu_m$. It follows from these remarks that if $x \in \mathbb{R}^n$ and m is the smallest integer such that $\Gamma_m = \{i \mid 1 \leq i \leq n\}$, then $m \leq n$ and $T_C^j(x) = \xi$ for all $j \geq m - 1$. In particular $T_C^j(x) = T_C^{j+1}(x)$ for all $j \geq n - 1$. A similar observation, using very different notation and techniques, seems to appear in chapter 27 of [7].

If one has further information about the matrix A , one can obtain sharper conclusions. For example, suppose that $a_{ij} = d_i > 0$ for $1 \leq j \leq n$ and $j \neq \sigma(i)$. If $x \in \mathbb{R}^n$ and $\mu_1 = \min_i x_i$ and $T_A(x) = y$, it is not hard to see that $\mu_1 \leq y_j \leq \mu_1 + d_j$ for $1 \leq j \leq n$ and $\min_j y_j = \mu_1$. If we define $Y = \{y \in \mathbb{R}^n : \mu_1 \leq y_j \leq \mu_1 + d_j \text{ for } 1 \leq j \leq n, \min_j y_j = \mu_1\}$, it follows that $T_A(Y) \subset Y$. If $y \in Y$, one can see that $T_A(y) = P(y)$, where P is as in remark 2.2. It follows that

$$T_C(T_A(x)) = T_A(x)$$

and one concludes that $T_C^{j+1} = T_C^j$ for all $j \geq 1$ in this case.

Remark 2.5. Elementary algebra tells us that the number p in theorem 2.2 is simply the order of some element σ of the permutation group S_n on n letters. If $G(n)$ denotes the maximal order of an element σ of S_n , a classical result of Landau (see [25] for references) asserts that

$$\log G(n) \sim \sqrt{n \log n}$$

Much more recently, Massias [26] has obtained effective bounds and proved that

$$\log G(n) \leq (1.053\ 13\ \dots) \sqrt{n \log n} \quad \text{for } n \geq 1$$

with equality occurring at $n = 1319\,766$. In [17, p 525] it is conjectured that if $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is non-expansive with respect to $\|\cdot\|_\infty$, then the period p of a periodic point x of f always satisfies $p \leq 2^n$, and in general such an upper bound, if true, is best possible. Our results here show that for our particular class of maps f (so $f = T_A$), much sharper results are true.

Acknowledgments

I would like to thank Robert Griffiths for informing me about his conjecture. Thanks are also due to John Mallet-Paret for some helpful early discussions about the questions treated here.

References

- [1] Akcoglu M A and Krengel U 1987 Nonlinear models of diffusion on a finite space *Prob. Theor. Rel. Fields* **76** 411–20
- [2] Blokhuis A and Wilbrink H A 1991 Alternative proof of Nussbaum's theorem on the size of a regular polygon in \mathbb{R}^n with the l_∞ -metric *Discrete Comput. Geom.* in press
- [3] Chou W and Duffin R J 1987 An additive eigenvalue problem of physics related to linear programming *Adv. Appl. Math.* **8** 486–98
- [4] Chou W and Griffiths R B 1986 Ground states of one-dimensional systems using effective potentials *Phys. Rev. B* **34** 6219–34
- [5] Crandall M G and Tartar L 1980 Some relations between non-expansive and order preserving mappings *Proc. Am. Math. Soc.* **78** 385–90
- [6] Cuninghame-Green R 1962 Describing industrial processes with interference and approximating their steady-state behaviour *Op. Res. Quart.* **13** 95–100
- [7] Cuninghame-Green R 1979 Minimax Algebra *Lecture Notes in Economics and Math. Systems* vol **166** (Berlin: Springer)
- [8] Floria L M and Griffiths R B 1989 Numerical procedure for solving a minimization eigenvalue problem *Numerische Math* **55** 565–74
- [9] Golitschek M V 1982 Optimal cycles in doubly weighted graphs and approximation of bivariate functions by univariate ones *Numerische Math* **39** 65–84
- [10] Graham A 1987 *Non-negative Matrices and Applicable Topics in Linear Algebra* (New York: Ellis Horwood)
- [11] Griffiths R B and Chou W 1986 Effective potentials, a new approach and new results for one-dimensional systems with competing length scales *Phys. Rev. Lett.* **56** 1929–31
- [12] Griffiths R B 1990 Personal letter to R D Nussbaum, May 15, 1990
- [13] Ishikawa S 1976 Fixed points and iteration of a non-expansive mapping in a Banach space *Proc. Am. Math. Soc.* **59** 65–71
- [14] Karp R M 1978 A characterization of the minimum cycle mean in a digraph, *Discrete Math.* **23** 309–11
- [15] Minc H 1988 *Non-negative Matrices* (New York: Wiley)
- [16] Misurewicz M Rigid sets in finite dimensional l_1 -spaces *preprint*
- [17] Nussbaum R D 1990 Omega limit sets of non-expansive maps: finiteness and cardinality estimates *Diff. Int. Eq.* **3** 523–40
- [18] Nussbaum R D 1988 Hilbert's projective metric and iterated nonlinear maps *Mem. Am. Math. Soc.* **391**
- [19] Nussbaum R D 1989 Hilbert's projective metric and iterated nonlinear maps, II *Mem. Am. Math. Soc.* **401**
- [20] Nussbaum R D Estimates of the period of periodic points for non-expansive operators *Israel J. Math.* in press
- [21] Scheutzwow M 1988 Periods of non-expansive operators on finite l_1 -spaces *Eur. J. Combinatorics* **9** 73–81
- [22] Seneta E 1981 *Non-negative Matrices and Markov Chains* (New York: Springer)

- [23] Sine R 1990 A nonlinear Perron–Frobenius theorem *Proc. Am. Math. Soc.* **109** 331–6
- [24] Weller D 1987 Hilbert’s metric, part metric and self-mappings of a cone *PhD Dissertation* University of Bremen
- [25] Miller W 1987 The maximum order of an element of a finite symmetric group *Am. Math. Monthly* **94** 497–506
- [26] Massias J-P 1984 Majoration explicite de l’ordre maximum d’un element du groupe symetrique *Ann. Fac. Sci. Toulouse* **VI** 269–81
- [27] Griffiths R B 1991 Frenkel–Kontorova models of commensurate–incommensurate phase transitions *Fundamental Problems in Statistical Mechanics* vol VII, ed H van Beijeren in press