A Note on Perturbations of Stochastic Matrices

B. Huppert[1] and W. Willems[2]

[1] Fachbereich Mathematik
Universität Mainz
55099 Mainz

[2] Fakultät für Mathematik
Otto-von-Guericke-Universität
39016 Madgeburg

HELMUT WIELANDT ZUM 90. GEBURTSTAG

Let A and B be stochastic matrices of the same type (n, n). It is natural to consider perturbations $A(t)$ of A of the form

$$A(t) = (1-t)A + tB \qquad (0 \leq t \leq 1).$$

Now, the following two questions come up immediately.

(1) When does $\lim_{k\to\infty} A(t)^k$ exist?

(2) How does $\lim_{k\to\infty} A(t)^k$ behave for small t?

As the existence of $\lim_{k\to\infty} A(t)^k$ depends only on the eigenvalues of $A(t)$, which are continously dependent on t, question (1) allows a positive answer (see section 1). The answer to question (2) is less simple. This is not surprising, for if

$$\lim_{k\to\infty} A(t)^k = \begin{pmatrix} v_1(t) \\ \vdots \\ v_n(t) \end{pmatrix},$$

then $v_i(t)A(t) = v_i(t)$ and it is well-known that eigenvectors do not behave well under perturbations.

1 The existence of $\lim_{k\to\infty} A(t)$.

Lemma 1.1 *Let A and B be stochastic matrices of the same type and let $A(t) = (1-t)A + tB$ for $0 \le t \le 1$. Furthermore let ϵ be a complex number with $|\epsilon| = 1$. Then the following two statements are equivalent.*

a) ϵ is an eigenvalue of $A(t_0)$ for some $0 < t_0 < 1$.

b) ϵ is an eigenvalue of $A(t)$ for all $0 \le t \le 1$.

Proof: We only have to prove that a) implies b) and we may assume that $\epsilon \ne 1$. By a theorem of Frobenius (see Wielandt [5], page 642) the eigenvalue ϵ is a primitive k-th root of unity for some $k \ne 1$. By the same theorem there exists a permutation matrix P such that

$$P^{-1}A(t_0)P = \begin{pmatrix} A_1(t_0) & & 0 \\ * & \ddots & 0 \\ * & * & A_m(t_0) \end{pmatrix}$$

with irreducible matrices $A_j(t_0)$. (Recall that a matrix is called *irreducible* if in the corresponding graph two vertices are connected by a path.) Now, if ϵ is an eigenvalue of $A_i(t_0)$, then we can furthermore assume that $A_i(t_0)$ is of the form

$$A_i(t_0) = \begin{pmatrix} 0 & C_1(t_0) & 0 & \dots & 0 \\ 0 & 0 & C_2(t_0) & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & C_{k-1}(t_0) \\ C_k(t_0) & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Let S_j denote the sum of the entries of $A_i(t_0)$ in its j-th row. By a theorem of Collatz and Wielandt (see [4], page 31) we have

$$\min_j S_j \le \mathrm{r}(A_i(t_0)) \le \max_j S_j \le 1$$

where $\mathrm{r}(A_i(t_0))$ denotes the spectral radius of $A_i(t_0)$. But

$$\mathrm{r}(A_i(t_0)) \ge |\epsilon| = 1.$$

Hence, again by the Collatz-Wielandt theorem, $S_j = 1$ for all j since $A_i(t_0)$ is irreducible. Therefore $A_i(t_0)$ is stochastic and

$$P^{-1}A(t_0)P = \begin{pmatrix} A_1(t_0) & & & & & & \\ & \ddots & & & & & \\ 0 & \dots & 0 & A_i(t_0) & 0 & \dots & 0 \\ & & & & \ddots & & \\ & * & & & & & A_m(t_0) \end{pmatrix}.$$

Now observe that an entry in $P^{-1}A(t_0)P$ is 0 if and only if the corresponding entries in $P^{-1}AP$ and $P^{-1}BP$ are both 0. Hence we obtain for t in $[0,1]$

$$P^{-1}A(t)P = \begin{pmatrix} A_1(t) & & & & & & \\ & \ddots & & & & & \\ 0 & \dots & 0 & A_i(t) & 0 & \dots & 0 \\ & & & & \ddots & & \\ & * & & & & & A_m(t) \end{pmatrix}$$

and

$$A_i(t) = \begin{pmatrix} 0 & C_1(t) & 0 & \dots & 0 \\ 0 & 0 & C_2(t) & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & C_{k-1}(t) \\ C_k(t) & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Furthermore $A_i(t)$ is stochastic and it is easy to see that ϵ is an eigenvalue of $A_i(t)$, hence of $A(t)$ for all $0 \le t \le 1$.

Theorem 1.2 *Suppose that A and B are stochastic matrices of the same type. If $\lim_{k\to\infty} A^k$ or $\lim_{k\to\infty} B^k$ exists, then*

$$\lim_{k\to\infty} ((1-t)A + tB)^k$$

exists for all $0 < t < 1$.

Proof: Assume that for some t_0 with $0 < t_0 < 1$ the limit $\lim_{k\to\infty} A(t_0)^k$ does not exist. Then $A(t_0)$ has an eigenvalue $\epsilon \neq 1$ with $|\epsilon| = 1$. By Lemma 1.1, ϵ is also an eigenvalue of $A(0) = A$ and $A(1) = B$. But then both $\lim_{k\to\infty} A^k$ and $\lim_{k\to\infty} B^k$ do not exist, a contradiction.

2 The behaviour of $\lim_{k\to\infty} A(t)^k$ for $AB = BA$.

Lemma 2.1 *Let A and B be stochastic matrices of the same type with $AB = BA$. If we put again $A(t) = (1-t)A + tB$, then for $0 < t < 1$ we have*

$$\ker(A(t) - E) = \ker(A - E) \cap \ker(B - E),$$

where E denotes the identity matrix.

Proof: Clearly,

$$\ker(A - E) \cap \ker(B - E) \subseteq \ker(A(t) - E).$$

As A and B commute, there exists a decomposition of the complex vector space V on which A and B act such that

$$V = V_1 \oplus \ldots \oplus V_r$$

where A and B induce lower triangular matrices of type

$$A_i = \begin{pmatrix} a_i & & \\ & \ddots & \\ * & & a_i \end{pmatrix} \quad \text{and} \quad B_i = \begin{pmatrix} b_i & & \\ & \ddots & \\ * & & b_i \end{pmatrix}$$

on V_i. Hence $(1-t)a_i + tb_i$ is the only eigenvalue of $A(t)$ on V_i. As $|a_i| \leq 1$ and $|b_i| \leq 1$, we have

$$(1-t)a_i + tb_i = 1 \quad \text{for} \quad 0 < t < 1$$

if and only if $a_i = b_i = 1$. But then $A_i = B_i = E$ as for stochastic matrices C we always have

$$\ker(C - E) = \ker(C - E)^2$$

(see [2], page 16, 2.6d)). This proves

$$\ker(A(t) - E) \subseteq \ker(A - E) \cap \ker(B - E)$$

and the proof is complete.

Theorem 2.2 *Let A and B be stochastic matrices of the same type with $AB = BA$. If we put*

$$Q(t) = \lim_{k\to\infty} \frac{1}{k} \sum_{j=0}^{k-1} A(t)^j,$$

then

$$Q(t) = Q(0)Q(1) \quad \text{for} \quad 0 < t < 1.$$

Proof: a) We set $R = Q(0)Q(1)$ and claim that

$$\operatorname{im} R = \ker(A - E) \cap \ker(B - E).$$

As A and B commute, so do the projections $Q(0)$ and $Q(1)$. Hence $R^2 = R$ is a projection as well. If $v = vA = vB$, then

$$v = vQ(0) = vQ(1),$$

and hence $v = vR$. Therefore we have

$$\ker(A - E) \cap \ker(B - E) \subseteq \ker(R - E) = \operatorname{im} R.$$

Suppose conversely that $v = wR \in \operatorname{im} R$. As $Q(0)A = Q(0)$, we obtain

$$vA = wQ(1)Q(0)A = wQ(1)Q(0) = v$$

and similary $vB = v$. Hence we have

$$\ker(R - E) \subseteq \ker(A - E) \cap \ker(B - E).$$

b) It is well-known (part of the ergodic theorem) that

$$\ker(A(t) - E) = \operatorname{im} Q(t).$$

Hence we obtain by a) and Lemma 2.1

$$\operatorname{im} R = \ker(A - E) \cap \ker(B - E) = \ker(A(t) - E) = \operatorname{im} Q(t)$$

for $0 < t < 1$. As $Q(t)$ and R commute, we have the $Q(t)$-invariant decomposition

$$V = \operatorname{im} R \oplus \ker R = \operatorname{im} Q(t) \oplus \ker R.$$

Thus $Q(t)$ induces on $\ker R$ a projection with all eigenvalues equal to 0. Therefore $\ker R = \ker Q(t)$ which finally shows that $R = Q(t)$.

Theorem 2.3 *Let A and B be stochastic matrices with $AB = BA$. Suppose that both $\lim_{k\to\infty} A^k = P_A$ and $\lim_{k\to\infty} B^k = P_B$ exist. Then we have*

$$\lim_{k\to\infty} A(t)^k = P_A P_B$$

for $0 < t < 1$.

Proof: By 1.2, the limit $\lim_{k\to\infty} A(t)^k$ exist. But then $\lim_{k\to\infty} A(t)^k = Q(t)$ and by 2.3,

$$\lim_{k\to\infty} A(t)^k = Q(0)Q(1) = P_A P_B.$$

Remarks 2.4 a) If we take

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

then $AB = BA$. One easily sees that

$$P_A = \lim_{k\to\infty} A^k = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad P_B = \lim_{k\to\infty} B^k = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Therefore $P_A \neq P_A P_B \neq P_B$. Hence in this case $P(t) = \lim_{k\to\infty} A(t)^k$ is constant for $0 < t < 1$ but discontinuous in $t = 0$ and $t = 1$.
b) Continuity of $P(t)$ in $t = 0$ in Theorem 2.3 requires $P_A = P_A P_B$. This is easily seen to be equivalent to

$$\ker(A - E) = \operatorname{im} P_A \subseteq \operatorname{im} P_B = \ker(B - E).$$

3 The behaviour of $\lim_{k\to\infty} A(t)^k$ for general A, B.

Lemma 3.1 *Let A and B be stochastic matrices of type (n, n). If A or B is irreducible, then*

$$A(t) = (1 - t)A + tB$$

is irreducible for $0 < t < 1$. In particular we have

$$\operatorname{rank}(A(t) - E) = n - 1 \quad \text{for } 0 < t < 1.$$

Proof: As irreducibility depends only on the zero entries in a matrix, the first statement is obvious. It is well-known that 1 is a simple eigenvalue of an irreducible stochastic matrix (see Wielandt [5]).

Theorem 3.2 *Let A and B be stochastic matrices of type (n, n). Suppose that $\lim_{k\to\infty} A^k$ exists and that $A(t)$ is irreducible for $0 < t < 1$.*

a) For $0 < t < 1$ *we have*

$$P(t) = \lim_{k\to\infty} A(t)^k = \begin{pmatrix} v(t) \\ \vdots \\ v(t) \end{pmatrix}$$

where the components of the vector $v(t)$ *are rational functions in* t *which are analytic in* $0 \le t \le 1$.

b) If $\operatorname{rank}(A - E) = n - 1$, *then* $P(t)$ *is continuous in* $t = 0$, *hence*

$$\lim_{k\to\infty} A^k = \lim_{t\to\ 0} P(t) = \begin{pmatrix} v(0) \\ \vdots \\ v(0) \end{pmatrix}.$$

Proof: By Theorem 1.2, $P(t) = \lim_{k\to\infty} A(t)^k$ exists for $0 \le t < 1$. As $\operatorname{rank}(A(t) - E) = n - 1$ for $0 < t < 1$, we have

$$P(t) = \begin{pmatrix} v(t) \\ \vdots \\ v(t) \end{pmatrix}$$

where $v(t)$ is the unique vector with $v(t)A(t) = v(t)$ and sum of components equal to 1. If $B(t)$ is the adjoint matrix of $A(t) - E$, then

$$B(t)(A(t) - E) = \det(A(t) - E)E = 0.$$

Now, since $\operatorname{rank}(A(t)-E) = n-1$, there is a row, say $z(t) = (z_1(t), \ldots, z_n(t))$ of $B(t)$ such that $z(t_0) \ne 0$ for some $0 < t_0 < 1$. As $A(t_0)$ is irreducible, we may assume that all components of the vector $z(t_0)$ which is annihilated by $A(t) - E$ are positive. In particular we have $\sum_{i=1}^n z_i(t_0) \ne 0$. Furthermore, since the $z_i(t)$ are polynomials in t, we may assume that $\sum_{i=1}^n z_i(t) \ne 0$ for an open intervall I containing t_0. If we now divide $z(t)$ by $\sum_{i=1}^n z_i(t)$, we obtain a vector $v(t) = (v_1(t), \ldots, v_n(t))$ with rational functions $v_i(t)$ in $t \in I$. Moreover,

$$v(t)A(t) = v(t) \text{ and } \sum_{i=1}^n v_i(t) = 1 \quad \text{for } t \in I.$$

But these identities must then hold for all $0 \le t \le 1$ except possibly for a finite number of poles of the $v_i(t)$.

Let $\tilde{t}$ be different from all of those poles and $0 \neq \tilde{t} \neq 1$. As $A(\tilde{t})$ is irreducible, $v(\tilde{t})$ is the unique solution of $v(\tilde{t})A(\tilde{t}) = v(\tilde{t})$ with $\sum_{i=1}^{n} v_i(\tilde{t}) = 1$. But then $v(\tilde{t})$ is a positive vector i.e. $0 < v_i(\tilde{t}) \leq 1$ for all i. This shows that the $v_i(t)$ have no poles in $0 \leq t \leq 1$.
b) Clearly, $v(0)A = v(0) \neq 0$ since $v(t)$ is analytic in $t = 0$. Furthermore, if rank $(A - E) = n - 1$, then $v(0)$ is the unique non-negative eigenvector of A for the eigenvalue 1 with sum of components equal to 1. Thus

$$\lim_{k \to \infty} A^k = \begin{pmatrix} v(0) \\ \vdots \\ v(0) \end{pmatrix}.$$

Remark 3.3 We consider again the case $AB = BA$ of Theorem 2.3, and suppose that rank $(A - E) = n - 1$. Then

$$P_A = \lim_{k \to \infty} A^k = \begin{pmatrix} v \\ \vdots \\ v \end{pmatrix}$$

where v is the unique non-negative vector with $vA = v$ and sum of components equal to 1. But then

$$(vB)A = vAB = vB$$

and vB is also non-negative with sum of components equal to 1. Thus we get $vB = v$ and therefore $P_A B = P_A$. From this we finally obtain

$$P_A P_B = \lim_{k \to \infty} P_A B^k = P_A.$$

Note that this is the continuity of $P(t)$ in $t = 0$ which follows immediately from 3.2 b) in case $A(t)$ is irreducible for $0 < t < 1$.

Now we turn to the most interesting case in which A has more than one absorbing state i.e. rank $(A - E) < n - 1$.

Theorem 3.4 *Suppose that*

$$A = \begin{pmatrix} E & 0 \\ C & D \end{pmatrix} \quad \textit{and} \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

are stochastic matrices of type (n,n) *and* E *is the identity matrix of type* (m,m) *with* $m>1$. *Furthermore, suppose that* $\lim_{k\to\infty} D^k = 0$ *which implies*

$$\lim_{k\to\infty} A^k = \begin{pmatrix} E & 0 \\ (E-D)^{-1}C & 0 \end{pmatrix}.$$

Finally, assume that $A(t) = (1-t)A + tB$ *is irreducible for* $0<t<1$. *Then we have*

a) $\lim_{k\to\infty} A(t)^k = \begin{pmatrix} v(t) & w(t) \\ \vdots & \vdots \\ v(t) & w(t) \end{pmatrix}$ *where* $w(0)=0$.

More precisely, it even holds that $\|w(t)\| \le 3t\|(E-D)^{-1}\|$ *where the matrix norm* $\|(c_{ij})\| = \max_i \sum_j |c_{ij}|$ *is used. In particular,*

$$\lim_{t\to\ 0}\lim_{k\to\infty} A(t)^k = \begin{pmatrix} v(0) & 0 \\ \vdots & \vdots \\ v(0) & 0 \end{pmatrix} \ne \lim_{k\to\infty} A^k$$

if A *has more than one absorbing state.*

b) There exists $\lim_{t\to\ 0} \frac{w(t)}{t} = w_1$, *and*

$$(v(0), w_1)\begin{pmatrix} B_{11} & B_{12} \\ C & D \end{pmatrix} = (v(0), w_1).$$

c) $w_1 \ne 0$. *Thus the effect of the perturbation in* $w(t)$ *is of first order in* t.

Proof: By Theorem 3.2, we have $\lim_{k\to\infty} A(t)^k = \begin{pmatrix} v(t) & w(t) \\ \vdots & \vdots \\ v(t) & w(t) \end{pmatrix}$ for all $0<t<1$ where the components of the vector $(v(t), w(t))$ are rational, and analytic in $0\le t\le 1$. The equation $(v(t), w(t)) = (v(t), w(t))A(t)$ implies

$$(1)\quad v(t) = (1-t)v(t) + tv(t)B_{11} + (1-t)w(t)C + tw(t)B_{21} \quad \text{and}$$

$$(2)\quad w(t) = tv(t)B_{12} + (1-t)w(t)D + tw(t)B_{22}.$$

From the second equation we get

$$\begin{aligned}
||w(t)(E-D)|| &= ||tv(t)B_{12} - tw(t)D + tw(t)B_{22}|| \\
&\leq t(||v(t)||\,||B_{12}|| + ||w(t)||\,||D|| + ||w(t)||\,||B_{22}|| \\
&\leq 3t.
\end{aligned}$$

Therefore

$$||w(t)|| = ||w(t)(E-D)(E-D)^{-1}|| \leq ||w(t)(E-D)||\,||(E-D)^{-1}|| \leq 3t||(E-D)^{-1}||.$$

In particular $w(0)=0$, and as w is rational, so $w_1 = \lim_{t\to\,0} \frac{w(t)}{t}$ exists.
b) Divide the equations (1) and (2) by t and let $t \to 0$. We get

$$\begin{aligned}
0 &= -v(0) + v(0)B_{11} + w_1 C \text{ and} \\
w_1 &= v(0)B_{12} + w_1 D.
\end{aligned}$$

This shows that

$$(v(0), w_1) = (v(0), w_1) \begin{pmatrix} B_{11} & B_{12} \\ C & D \end{pmatrix}.$$

c) Suppose that $w_1 = 0$. Then the equations in b) reduce to

$$v(0)B_{11} = v(0) \text{ and } v(0)B_{12} = 0.$$

Hence we obtain

$$(v(0), 0)A(t) = (v(0), 0) \text{ for } 0 \leq t \leq 1.$$

As the non-negative eigenvector of the irreducible matrix $A(t)$ (for $0 < t < 1$) has only non-zero components (see Wielandt, [5]), we have a contradiction. Thus $w_1 \neq 0$.

Remarks 3.5 a) We consider again the situation in Theorem 3.4. By Theorem 3.2, the vectors $v(t)$ and $w(t)$ are analytic in $[0,1]$. Thus we may write $v(t) = \sum_{j=0}^{\infty} v_j t^j$ and $w(t) = \sum_{j=0}^{\infty} w_j t^j$. As $(v(t), w(t))$ has sum of components equal to 1, we obtain that v_0 has sum of components equal to 1 and (v_j, w_j) has sum of components equal to 0 for $j > 0$. To be brief in the following we denote by $\mathrm{s}(v)$ the sum of components of the vector v. From

$$(v(t), w(t)) = (v(t), w(t))A(t)$$

we get by comparing coefficients

$$(1) \qquad (v_0, w_1)\left(\begin{pmatrix} B_{11} & B_{12} \\ C & D \end{pmatrix} - E\right) = 0$$

and for $j > 1$

$$(2)\ (v_{j-1}, w_j)\left(\begin{pmatrix} B_{11} & B_{12} \\ C & D \end{pmatrix} - E\right) = (w_{j-1}(C - B_{21}), w_{j-1}(D - B_{22})).$$

Now we assume that

$$\operatorname{rank}\left(\begin{pmatrix} B_{11} & B_{12} \\ C & D \end{pmatrix} - E\right) = n - 1.$$

Then (v_0, w_1) is uniquely determined by (1) and the condition $\mathrm{s}(v_0) = 1$. Obviously,

$$\mathrm{s}((0, w_{j-1})(A - B)) = \mathrm{s}((w_{j-1}(C - B_{21}), w_{j-1}(D - B_{22})) = 0.$$

Hence $(w_{j-1}(C - B_{21}), w_{j-1}(D - B_{22}))$ is contained in the space

$$U = \{ (x_1, \ldots, x_n) \mid \sum_{j=1}^{n} x_j = 0\}.$$

For the same reason we have

$$\operatorname{im}\left(\begin{pmatrix} B_{11} & B_{12} \\ C & D \end{pmatrix} - E\right) \subseteq U.$$

As

$$\operatorname{rank}\left(\begin{pmatrix} B_{11} & B_{12} \\ C & D \end{pmatrix} - E\right) = n - 1,$$

we obtain

$$(w_{j-1}(C - B_{21}), w_{j-1}(D - B_{22})) \subseteq U = \operatorname{im}\left(\begin{pmatrix} B_{11} & B_{12} \\ C & D \end{pmatrix} - E\right).$$

Hence we can calculate (v_{j-1}, w_j) uniquely up to a scalar multiple of (v_0, w_1) from the equation (2). As we already know w_{j-1} by the earlier step, we also

know $\mathrm{s}(v_{j-1}) = -\mathrm{s}(w_{j-1})$. This determines (v_{j-1}, w_j) uniquely.
b) Unfortunately such an algorithm does not seem to be available if

$$\operatorname{rank}\left(\begin{pmatrix} B_{11} & B_{12} \\ C & D \end{pmatrix} - E\right) < n-1.$$

Consider for example

$$A = \begin{pmatrix} E & 0 \\ \frac{1}{2}E & \frac{1}{2}E \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} \frac{1}{2}E & \frac{1}{2}E \\ F & F \end{pmatrix}$$

where all matrices E, F are of the same type (n, n) and $2F$ is stochastic. To garantee the irreducibility of $A(t)$ for $0 < t < 1$ we may assume that F is strictly positive i.e. all entries are positive. From

$$(v(t), w(t)) = (v(t), w(t))A(t)$$

we easily compute $\frac{1}{2}v(t) = v(t)F$ and $w(t) = tv(t)$. If we choose v with sum of components equal to 1 such that $\frac{1}{2}v = vF$, we finally get

$$\lim_{k\to\infty} A(t)^k = \begin{pmatrix} \frac{v}{1+t} & \frac{tv}{1+t} \\ \vdots & \vdots \\ \frac{v}{1+t} & \frac{tv}{1+t} \end{pmatrix}$$

for $0 < t < 1$. Therefore, if $n \geq 2$, then

$$\lim_{t\to 0}\lim_{k\to\infty} A(t)^k = \begin{pmatrix} v & 0 \\ \vdots & \vdots \\ v & 0 \end{pmatrix} \neq \lim_{k\to\infty} A^k = \begin{pmatrix} E & 0 \\ E & 0 \end{pmatrix}.$$

But the equation

$$(v(0), w_1) = (v(0), w_1)\begin{pmatrix} \frac{1}{2}E & \frac{1}{2}E \\ \frac{1}{2}E & \frac{1}{2}E \end{pmatrix}$$

only implies $v(0) = w_1$, hence does not allow the calculation of $v(0)$ and w_1.

It seems that in general there is no way to calculate $v(0)$ without determining $v(t)$ for all small t.

Examples 3.6 a) Let

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & & & \\ a_{21} & a_{22} & a_{23} & 0 & & & \\ 0 & a_{32} & a_{33} & a_{34} & & & \\ & & & & & & \\ & & & & a_{n-1,n-2} & a_{n-1,n-1} & a_{n-1,n} \\ & & & & 0 & 0 & 1 \end{pmatrix}$$

and let

$$B = \begin{pmatrix} b_{11} & b_{12} & 0 & 0 & & & \\ b_{21} & b_{22} & b_{23} & 0 & & & \\ 0 & b_{32} & b_{33} & b_{34} & & & \\ & & & & & & \\ & & & & b_{n-1,n-2} & b_{n-1,n-1} & b_{n-1,n} \\ & & & & 0 & b_{n,n-1} & b_{nn} \end{pmatrix}$$

be stochastic Jacobi matrices of type (n, n). We assume that

$$a_{21}a_{32}\ldots a_{n-1,n-2}a_{23}a_{34}\ldots a_{n-1,n} \neq 0 \neq b_{12}b_{n,n-1}.$$

Thus

$$A(t) = (1-t)A + tB$$

is irreducible for $0 < t < 1$. By recursion we find the eigenvector $v(t)$ with $v(t) = v(t)A(t)$, namely $v(t) = (v_1(t), \ldots, v_n(t))$ where

$$\frac{v_j(t)}{v_1(t)} = \frac{tb_{12}}{(1-t)a_{21} + tb_{21}} \cdot \frac{(1-t)a_{23} + tb_{23}}{(1-t)a_{32} + tb_{32}} \cdots \frac{(1-t)a_{j-1,j} + tb_{j-1,j}}{(1-t)a_{j+1,j} + tb_{j+1,j}}$$

for $2 \leq j < n$ and

$$\frac{v_n(t)}{v_1(t)} = \frac{tb_{12}}{(1-t)a_{21} + tb_{21}} \cdots \frac{(1-t)a_{n-1,n} + tb_{n-1,n}}{tb_{n,n-1}}.$$

Therefore

$$\lim_{t\to 0} v(t) = (v_1(0), 0, \ldots, 0, v_n(0))$$

where

$$v_n(0) = v_1(0)\frac{a_{23}a_{34}\ldots a_{n-1,n}}{a_{21}a_{32}\ldots a_{n-1,n-2}} \cdot \frac{b_{12}}{b_{n,n-1}} =: v_1(0)c.$$

Naturally, we also have $v_1(0) + v_n(0) = 1$ which determines $v_1(0) = \frac{1}{1+c}$. Thus we obtain

$$\lim_{t\to 0}\lim_{k\to\infty} A(t)^k = \begin{pmatrix} \frac{1}{1+c} & 0 & \dots & 0 & \frac{c}{1+c} \\ \vdots & & & & \vdots \\ \frac{1}{1+c} & 0 & \dots & 0 & \frac{c}{1+c} \end{pmatrix},$$

but

$$\lim_{k\to\infty} A^k = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ d_2 & 0 & \dots & 0 & 1-d_2 \\ \vdots & \vdots & & \vdots & \vdots \\ d_{n-1} & 0 & \dots & 0 & 1-d_{n-1} \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

for some d_i.

b) A special case of Example a) is the Moran model with two alleles, say a and b (see Ewens, page 84-89). The transition matrix is a Jacobi matrix given by $A = (a_{ij})$ $(i, j = 0, \dots, n)$ with

$$a_{i,i-1} = a_{i,i+1} = \frac{i(n-i)}{n^2} \text{ and } a_{ii} = 1 - a_{i,i-1} - a_{i,i+1}.$$

If a mutation from a to b with rate t and from b to a with rate ct (c fixed) is allowed, the transition matrix becomes $A(t) = (a_{ij}(t))$ with

$$\begin{aligned} a_{i,i-1}(t) &= \tfrac{i(n-i)(1-ct)+ti^2}{n^2} \\ a_{i,i+1}(t) &= \tfrac{i(n-i)(1-t)+ct(n-i)^2}{n^2} \\ a_{ii}(t) &= 1 - a_{i,i-1}(t) - a_{i,i+1}(t) \end{aligned}$$

Here we have $A(t) = (1-t)A + tB$ where B is a Jacobi matrix $B = (b_{ij})$ with

$$\begin{aligned} b_{i,i-1} &= \tfrac{i(n-i)(1-c)+i^2}{n^2} \\ b_{i,i+1} &= c\tfrac{(n-i)^2}{n^2} \\ b_{ii} &= 1 - b_{i,i-1} - b_{i,i+1}. \end{aligned}$$

By part a) we obtain

$$\lim_{t\to 0}\lim_{k\to\infty} A(t)^k = \begin{pmatrix} \frac{1}{1+c} & 0 & \dots & 0 & \frac{c}{1+c} \\ \vdots & & & & \vdots \\ \frac{1}{1+c} & 0 & \dots & 0 & \frac{c}{1+c} \end{pmatrix}.$$

Let $\rho(A(t)) = \max\{\,|\lambda| \mid |\lambda| < 1, \lambda$ an eigenvalue of $A(t)\}$. Note that $\rho(A(t))$ gives us some information about the speed of convergence of $A(t)^k$ for $k \to \infty$.

First we deal with the case that $t = 0$, i.e. $A(0) = A$. A simple calculation shows that

$$Av = (1 - \frac{2}{n^2})v \quad \text{for} \quad v = \begin{pmatrix} 0 \\ 1(n-1) \\ 2(n-2) \\ \vdots \\ 1(n-1) \\ 0 \end{pmatrix}.$$

Thus,

$$Av = \begin{pmatrix} 1 & 0 & 0 \\ * & C & * \\ 0 & 0 & 1 \end{pmatrix} v = (1 - \frac{2}{n^2})v$$

implies that

$$Cw = (1 - \frac{2}{n^2})w \quad \text{for} \quad w = \begin{pmatrix} 1(n-1) \\ 2(n-2) \\ \vdots \\ 1(n-1) \end{pmatrix}.$$

As C is irreducible and w strictly positive, a theorem of Perron and Frobenius (see [3]) says that $1 - \frac{2}{n^2}$ is the spectral radius of C. This proves that $\rho(A) = 1 - \frac{2}{n^2}$.

For t arbitrary $(0 \leq t \leq 1)$ a tediuos calculation shows that

$$A(t)u = \left(1 - \frac{(1+c)t)}{n}\right) u \quad \text{for} \quad u = \begin{pmatrix} 0 \\ 1 \\ 2 \\ \vdots \\ n \end{pmatrix} - \frac{cn}{1+c}\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Thus $\rho(A(t)) \geq 1 - \frac{(1+c)t}{n}$. Even equality holds here (see [1], page 88). Note that for n large the perturbated system $(0 < t < 1)$ converges faster than the non-perturbated one $(t = 0)$.

References

[1] W.J. Ewens, Mathematical Population Genetics, Springer-Verlag 1979.

[2] F.J. Fritz, B. Huppert and W. Willems, Stochastische Matrizen, Springer-Verlag 1979.

[3] G. Frobenius, Über Matrizen aus nicht negativen Elementen. S.-B. Deutsch. Akad. Wiss. Berlin, Math.-Nat. Kl. (1912), 456-477.

[4] R.S. Varga, Matrix Iterative Analysis, Prentice Hall 1962.

[5] H. Wielandt, Unzerlegbare, nichtnegative Matrizen, Math.Z. 52 (1950), 642-648.