Report S 299

On K.L. Chung's problem of imbedding a time-discrete Markov chain in a time-continuous one for finitely many states I

by

J.Th. Runnenburg

April 1962

Call a stochastic matrix (i.e. a square matrix of order n with elements $p_{jk} \geq 0$ for all $j,k \in \{1,2,\ldots,n\}$ and $\sum_{k=1}^{n} p_{jk}=1$ for all $j \in \{1,2,\ldots,n\}$) a P-matrix. Then the following problem (further called Chung's problem) is a specialization of a problem posed in CHUNG [1958] and CHUNG [1960] .

<u>Chung's problem</u>: Find conditions in order that to a given P-matrix P_1 there exists a matrix function P(t) such that P(t) is a P-matrix for each $t \geq 0$, for which [1])

$$
(1) \qquad \begin{cases} \lim\limits_{t \downarrow 0} P(t)=I, \\ P(s+t)=P(s)P(t) \quad \text{for all real } s,t > 0, \\ P(1)=P_1. \end{cases}
$$

For n=2 Chung's problem has been solved by FRÉCHET 1952 and BELLMAN 1960. In my thesis (RUNNENBURG 1960) the case n=3 was considered.

It is well-known (cf. DOOB [1953] , BIRKHOFF and VARGA [1958]), that if P(t) exists, then

$$
(2) \qquad P(t)=e^{Q_1 t} \qquad \text{for all real } t \geq 0,
$$

where Q_1 is a Q-matrix (i.e. a square matrix of order n with elements q_{jk} satisfying $q_{jk} \geq 0$ if $j \neq k$ for all $j,k \in \{1,2,\ldots,n\}$ and $\sum_{k=1}^{n} q_{jk}=0$ for all $j \in \{1,2,\ldots,n\}$).

Furthermore, if $\mu_1, \mu_2, \ldots, \mu_n$ are the eigenvalues of Q_1, then the eigenvalues $\lambda_1(t), \lambda_2(t), \ldots, \lambda_n(t)$ of P(t) are given by

$$
(3) \qquad \lambda_j(t)=e^{\mu_j t} \qquad \text{for all real } t \geq 0 \text{ and all } j.
$$

----

1) I is the $n \times n$ identity matrix.

According to GANTMACHER [1959] , in 1938 Kolmogorov posed the problem: Characterize the complex numbers z with $|z| \leq 1$ which can occur as eigenvalues of an n^{th} order stochastic matrix. This problem was partly solved in DMITRIEV and DYNKIN [1946] and definitely in KARPELEWITSCH [1951] . Making use of their results, the next theorem can be proved.

<u>Theorem</u>: The eigenvalues $\mu_1, \mu_2, \ldots, \mu_n$ of a Q-matrix (of order $n \geq 3$) satisfy

$$(4) \qquad (\tfrac{1}{2}+\tfrac{1}{n})\pi \leq \arg \mu_j \leq (\tfrac{3}{2}-\tfrac{1}{n})\pi \quad \text{for } 1 \leq j \leq n;$$

the only Q-matrices Q_1^* (with elements q_{jk}^*) with at least one $\mu_j \neq 0$ on the boundary of this region are given (after a suitable renumbering of states, i.e. rows and columns at the same time) by

$$(5) \qquad q_{jk}^* = \begin{cases} -\alpha & \text{for } j \equiv k \pmod{n}, \\ \alpha & \text{for } j \equiv k-1 \pmod{n}, \\ 0 & \text{otherwise}, \end{cases}$$

where α is an arbitrary positive number.

<u>Remark 1</u>: From this theorem we conclude, that for $n \geq 3$ all eigenvalues $\lambda_j(t)$ of $P(t)=e^{Q_1 t}$ satisfy

$$(6) \qquad \lambda_j(t) \in H_n,$$

where H_n is a heart-shaped region in the complex plane, contained in the unit circle and symmetric with respect to the real axis, with the curve

$$(7) \qquad \exp(-1+\cos\frac{2\pi}{n} + i \sin\frac{2\pi}{n})t \quad \left(\text{where } 0 \leq t \leq \frac{\pi}{\sin\frac{2\pi}{n}}\right)$$

as boundary in the region $\operatorname{Im} z \geq 0$. Hence Chung's problem can only be solved for those P_1 for which all eigenvalues belong to the region H_n. The only matrices e^{Q_1} with a Q-matrix Q_1 for which at least one eigenvalue $\lambda \neq 1$ lies on the boundary curve of the region H_n are given by

(8) $$P_1^* = e^{Q_1^*}.$$

Proof of the theorem: If P_1 is a P-matrix, then $Q_1=P_1-I$ is a Q-matrix; if Q_1 is a Q-matrix, then $P_{1,\beta}=I+\beta Q_1$ (where $\beta>0$ is chosen in such a way that $0\le 1+\beta q_{jj}\le 1$ for $1\le j\le n$, e.g. $\beta=-\left(\min\limits_{1\le j\le n} q_{jj}\right)^{-1}$ if some $q_{jj}<0$) is a P-matrix. Hence if for an arbitrary Q-matrix Q_1 we have

(9) $$\det(Q_1-\mu I)=0,$$

then

(10) $$\det(P_{1,\beta}-(\beta\mu+1)I)=0.$$

Now Karpelewitsch has shown that the roots of the characteristic equation

(11) $$\det(P_1-\lambda I)=0$$

for an arbitrary P-matrix P_1, satisfy

(12) $$\left(\frac{1}{2}+\frac{1}{n}\right)\pi \le \arg(\lambda-1) \le \left(\frac{3}{2}-\frac{1}{n}\right)\pi$$

for $n\ge 3$, and so with $\lambda=\beta\mu+1$, because of $\beta>0$ we have

(13) $$\left(\frac{1}{2}+\frac{1}{n}\right)\pi \le \arg\mu \le \left(\frac{3}{2}-\frac{1}{n}\right)\pi .$$

If Q_1^* has an eigenvalue μ_j for which $\arg\mu_j=\left(\frac{1}{2}+\frac{1}{n}\right)\pi$ with $\mu_j\ne 0$, i.e. $\mu_j=\gamma\left(e^{\frac{2\pi i}{n}}-1\right)$ with a $\gamma>0$. then $P_{1,\beta}^*$ has an eigenvalue λ_j with $\arg(\lambda_j-1)=\left(\frac{1}{2}+\frac{1}{n}\right)\pi$ and $\lambda_j-1\ne 0$. As Dmitriev and Dynkin have shown, in that case the elements p_{jk}^* of $P_{1,\beta}^*$ satisfy (after a suitable renumbering of the rows and columns at the same time)

(14) $$p_{jk}^* = \begin{cases} 1-\alpha_j' & \text{for } j\equiv k \pmod n, \\ \alpha_j' & \text{for } j\equiv k-1 \pmod n, \\ 0 & \text{otherwise,} \end{cases}$$

where $0 \leq \alpha_j' \leq 1$ for each $j \in \{1,2,\ldots,n\}$ .Therefore the elements q_{jk}^* of Q_1^* are given by

$$(15) \qquad q_{jk}^* = \begin{cases} -\alpha_j & \text{for } j \equiv k \pmod n \\ \alpha_j & \text{for } j \equiv k-1 \pmod n \\ 0 & \text{otherwise} \end{cases}$$

where we have written α_j instead of $\frac{\alpha_j'}{\beta}$, and so $\alpha_j \geq 0$ for each $j \in \{1,2,\ldots,n\}$.

For the characteristic equation of Q_1^* we obtain

$$(16) \qquad (\alpha_1+\mu)(\alpha_2+\mu)\ldots(\alpha_n+\mu) = \alpha_1 \alpha_2 \ldots \alpha_n .$$

In order that $\mu = \gamma(e^{\frac{2\pi i}{n}} -1)$, with a constant $\gamma > 0$, is a root of this equation, all α_j must be positive. We shall prove that

$$(17) \qquad \alpha_j = \gamma \quad \text{for all} \quad j \in \{1,2,\ldots,n\} ,$$

for which it is sufficient to consider $\gamma = 1$.

If we introduce the finite positive numbers $\theta_j = \frac{1}{\alpha_j}$ in (16) and substitute $e^{\frac{2\pi i}{n}} -1$ for μ , we obtain

$$(18) \qquad \left\{\theta_1 e^{\frac{2\pi i}{n}} +(1-\theta_1)\right\} \left\{ \theta_2 e^{\frac{2\pi i}{n}}+(1-\theta_2)\right\} \ldots \left\{\theta_n e^{\frac{2\pi i}{n}}+(1-\theta_n)\right\} =1.$$

Consider a line in the complex plane which does not pass through the origin. Let $r(\varphi)$ be the absolute value of the complex number on this line with argument φ . Then $\log r(\varphi)$ is a strictly convex analytic function of φ for finite $r(\varphi)$. Hence if $\varphi_1 \neq \varphi_2$, we have for any $p_1, p_2 > 0$ with $p_1+p_2=1$

$$(19) \qquad p_1 \log r(\varphi_1)+p_2 \log r(\varphi_2) > \log r(p_1\varphi_1+p_2\varphi_2).$$

Therefore

$$(20) \qquad r(\varphi_1)\, r(\varphi_2)\ldots r(\varphi_n) > r\left(\frac{\varphi_1+\varphi_2+\ldots+\varphi_n}{n}\right)^n ,$$

unless $\varphi_1 = \varphi_2 = \ldots = \varphi_n$.

Now consider for $n \geq 3$ the points $\theta_j . e^{\frac{2\pi i}{n}} + (1-\theta_j).1$ for finite $\theta_j > 0$, which lie on the line passing through 1 and $e^{\frac{2\pi i}{n}}$. Then $0 < \varphi_j < (\frac{1}{2} + \frac{1}{n})\pi$. Under the restriction $\varphi_1 + \varphi_2 + \ldots + \varphi_n \equiv 0 \pmod{2\pi}$, the product $r(\varphi_1)r(\varphi_2)\ldots r(\varphi_n)$ assumes its smallest value for $\varphi_1 = \varphi_2 = \ldots = \varphi_n = \frac{2\pi}{n}$. Any other choice of values for the φ_j leads to a larger value for the product. Hence if we introduce polar coordinates in (18) by writing

$$\theta_j \, e^{\frac{2\pi i}{n}} + 1 - \theta_j = r(\varphi_j) e^{i\varphi_j} \, , \tag{21}$$

we find that under the condition "all θ_j are finite positive numbers" the left-hand side of (18) has smallest positive value 1, where the value 1 is only obtained if $\varphi_j = \frac{2\pi}{n}$ for all j. Therefore, if equation (18) holds, we must have $\theta_j = 1$ for all j or $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 1$.

<u>Remark 2</u>: For n=3 one may easily verify that there exist P-matrices P_1 for which Chung's problem does <u>not</u> have a solution, although P_1 has all eigenvalues inside H_3. If P_1 is a circulant (i.e. a square matrix of order n with elements $p_{jk} = p_{k-j}$, where k-j is taken modulo n), then for n=3 a circulant Q_1 can be found which is a Q-matrix and for which $P_1 = e^{Q_1}$ holds, on condition that all eigenvalues of P_1 are inside H_3.

<u>References</u>

Bellman [1960] ,R, Introduction to Matrix Analysis, Mc Graw-Hill, New York, page 268.

Birkhoff and Varga [1958] ,G and R.S.,Reactor Criticality and Nonnegative Matrices, Journ. of Ind. and Appl.Math. <u>6</u> (1958)354-377.

Chung [1958],K.L., Continuous parameter Markov chains, Proc.Int.Congr.Math.Edinburgh, page 516.

Chung [1960], K.L., Markov Chains with Stationary Transition Probabilities, Springer, Berlin, page 203.

Dmitriev and Dynkin |1946|,N. and E., On characteristic roots of stochastic matrices, Isvestija, ser. mat., <u>10</u> (1946) 167-194.

Doob [1953],J.L., Stochastic Processes, John Wiley, New York, Chapter VI, § 1.

Fréchet [1952],M., Traité du Calcul des Probabilités et ses Applications, Tome I, Fascicule III, Second Livre, Méthode des Fonctions Arbitraires, Gauthier-Villars, Paris, page 255.

Gantmacher [1959],F.R., Matrizenrechnung II, Deutscher Verlag der Wissenschaften, Berlin, page 76,

Karpelewitsch [1951],F.I., On the characteristic roots of a matrix with nonnegative elements, Isvestija, ser.mat. <u>15</u> (1951) 361-383.

Runnenburg [1960],J.Th. On the Use of Markov Processes in One-Server Waiting-Time Problems and Renewal Theory, Thesis, Poortpers, Amsterdam, Stellingen XII and XIII.

Report S 299

On K.L. Chung's problem of imbedding a time-discrete Markov chain in a time-continuous one for finitely many states II *)

by

C.L. Scheffer

April 1962

*) Part I of this report was written before we knew of KINGMAN [5]. The problem is really due to G. ELFVING [6].

1 The purpose of this note is to show that one of the results obtained in the previous one [1], can be deduced without using the work of DMITRIEV and DYNKIN [2] and KARPELEWITSCH [3]. For a statement of the problem we refer to [1].

We use the fact that a Markov-chain on a finite set consisting of n states $A_1, A_2, \ldots, A_n$, induces an abelian semi-group G of real linear transformations on an R_n: with the Markov matrix (p_{jk}) (where $p_{jk} = P\{\underline{x}_{n+1} = A_k \mid \underline{x}_n = A_j\}$) corresponds the linear transformation π defined by

$$\pi e_j = (p_{j1}, p_{j2}, \ldots, p_{jn}),$$

where e_j is the j'th unit vector.

A necessary and sufficient condition that the matrix of a given linear transformation φ is a Markov matrix is

$$(1) \qquad \varphi W \subset W,$$

where W is the set of all probability distributions on $\{A_1, A_2, \ldots, A_n\}$

$$(2) \qquad W = \{x \mid x = (\xi_1, \xi_2, \ldots, \xi_n);\ \forall_j\ \xi_j \geq 0;\ \sum_{j=1}^{n} \xi_j = 1\}.$$

2 Now let G be a continuous one-parameter abelian semi-group of linear transformations satisfying (1) (ι is the identical transformation on R_n)

$$G = \{\varphi_t \mid t \in (0,\infty);\ \forall_t^{(0,\infty)}\ \varphi_t W \subset W;\ \forall_t^{(0,\infty)} \forall_s^{(0,\infty)}\ \varphi_t \varphi_s = \varphi_{t+s};\ \lim_{t \downarrow 0} \varphi_t = \iota\}.$$

It is well-known (cf e.g. FRÉCHET [4]) that under these conditions the eigenvalues of φ_t are $1, \lambda_1^t, \lambda_2^t, \ldots, \lambda_m^t$ ($m \leq n-1$), where $1, \lambda_1, \lambda_2, \ldots, \lambda_m$ are eigenvalues of φ_1, for some choice of the arguments of $\lambda_1, \ldots, \lambda_m$. Moreover there is at least one vector $p \in W$, which is invariant under G.

Let $\lambda = \varrho e^{i\psi}$ (ϱ real and ≥ 0) be an eigenvalue of φ_1. We suppose that ψ is chosen in such a way that $\varrho^t e^{it\psi}$ is, for all $t > 0$, an eigenvalue of φ_t. Moreover we suppose for the moment $\psi \neq 0$; $\varrho \neq 0$ and $n \geq 3$. (The inequality (20) which we want to establish is trivially satisfied for $\psi=0$ or $\varsigma=0$).

Let

$$z_o = x_o+iy_o,$$

where x_o and y_o are real vectors, be an eigenvector of φ_1 belonging to the eigenvalue λ. As $\psi\neq 0$, it follows that x_o and y_o are independent. Then the two-dimensional subspace V spanned by x_o and y_o

$$V = \{ u \mid u=\alpha x_o+\beta y_o;\ \alpha,\beta \ \text{real}\}$$

is invariant under G, and the movement induced by G in V is easily seen to be described by the equations

$$(3)\qquad \begin{cases} \varphi_t\, x_o = \varsigma^t(x_o \cos t\psi - y_o \sin t\psi), \\ \varphi_t\, y_o = \varsigma^t(x_o \sin t\psi + y_o \cos t\psi). \end{cases}$$

Introducing

$$W_o = \{ x \mid x\in W;\ \exists_j\ \xi_j = 0 \},$$

$$U = W \cap (p+V),$$

$$U_o = W_o \cap (p+V),$$

the condition (1) together with the invariance op p and V yields

$$(4)\qquad \forall_t^{(0,\infty)}\ \varphi_t\, U \subset U .$$

As U is a closed convex polygon[1] in the plane p+V and

----

1) It might seem that this statement is false for n=3. However in that case our hypothesis $\psi\neq 0$ entails that 1, $\varsigma e^{it\psi}$ and $\varsigma e^{-it\psi}$ are the only eigenvalues of φ_t (and hence of the adjoint transformation).
Therefore $\{ x \mid \xi_1+\xi_2+\xi_3=0\}$ is the only possible real invariant two-dimensional subspace and thus $V=\{ x \mid \xi_1+\xi_2+\xi_3=0\}$, implying $W\subset p+V$, whence $U=W$ and $U_o=W_o$.

U_o its boundary the condition (4) is satisfied if and only if

(5) $$\forall_t^{(0,\infty)} \quad \varphi_t U_o \subset U$$

is satisfied.

To derive a geometric condition equivalent with (5), consider a vector $w \in U_o$. The curve $\{\varphi_t w | t > 0\}$ must, on account of (5), be in U. Hence a vector v_1 tangent to $\{\varphi_t w | t \geqq 0\}$ in w must point from w to the interior of U. To make this condition more precise consider a vector v_2 pointing from w along U_o in such a way that v_1 lies inside the angle[1] formed by p-w and v_2.

If now we define

(6) $$\Theta_w = \sphericalangle(p-w, v_1); \; \Theta_{o,w} = \sphericalangle(p-w, v_2),$$

then it is easily seen that (5) is satisfied if and only if

(7) $$\forall_{w \in U_o} \; \Theta_w \leqq \Theta_{o,w}$$

is true.

3 It does not seem possible to deduce useful results directly from (7), as Θ_w is in general a complicated expression depending on w. The only case in which Θ_w is independent of w obtains when $|x_o| = |y_o|$ and the inner product $(x_o, y_o) = 0$, which will be shown below. The general case can be reduced to this case by means of the following considerations:

Let σ be a real non-singular linear transformation on R_n. Then the semi-group

$$G^* = \sigma G \sigma^{-1}$$

----

1) There is a slight difficulty here on account of the fact that p may be in U_o. This case does not occur if $\psi \neq 0$, $\varsigma \neq 0$: On account of (3) any vector w-p performs a complete rotation around p in the time interval $(0, 2\pi/\psi]$. Hence (5) and $p \in U_o$ are inconsistent.

consists of those linear transformations φ_t^* for which

$$\varphi_t^* = \sigma \varphi_t \sigma^{-1} .$$

These φ_t^* have the same eigenvalues as φ_t and eigenvectors

$$z^* = \sigma z, \text{ with } z^* = x^* + iy^* \ ; \ x^* = \sigma x; \ y^* = \sigma y,$$

where $z=x+iy$ is an eigenvector of φ_t.

Application of the transformation σ to $W, p, x_o, y_o, V, U, U_o, v_1, v_2$ yields $W^*, p^*, x_o^*, y_o^*, V^*, U^*, U_o^*, v_1^*, v_2^*$, i.e. $W^* = \sigma W$, etc.

We can now repeat the argument of 2 and arrive at equivalent conclusions. In particular, if we define

$$(6^*) \qquad \theta_w^* = \measuredangle (p^* - w, v_1^*); \ \theta_{o,w}^* = \measuredangle (p^* - w, v_2^*),$$

then

$$(7^*) \qquad \forall_{w \in U_o^*} \ \theta_w^* = \theta_{o,w}^*$$

is equivalent with (7).

In view of the remark at the beginning of this section, we choose σ in such a way that

$$(8) \qquad |x_o^*| = |y_o^*| \ \text{ and } (x_o^*, y_o^*) = 0.$$

Using (8) it is easy to calculate θ_w^*. For an arbitrary $w \in p^* + V^*$ we have

$$(9) \qquad w = p^* + \alpha(x_o^* \sin \gamma + y_o^* \cos \gamma) \quad (\alpha, \gamma \ \text{ real}).$$

Consider $\{\varphi_t \, w \mid t \geq 0\}$. A vector tangent to this curve in w is given by

$$(10) \qquad v_1^* = \lim_{h \downarrow 0} \frac{\varphi_h^* w - w}{h} .$$

Substitution of (9) and

$$(3^*) \qquad \begin{cases} \varphi_t^* x_o^* = \rho^t (x_o^* \cos t\psi - y_o^* \sin t\psi) \\ \varphi_t^* y_o^* = \rho^t (x_o^* \sin t\psi + y_o^* \cos t\psi) \end{cases}$$

in (10) gives after some calculation

$$(11)\qquad v_1^* = -\alpha A \left[x_o^* \sin(\gamma-\chi)+y_o^* \cos(\gamma-\chi)\right] ,$$

where

$$(12)\qquad A^2=\psi^2+\log^2\varrho \ ; \ 0<\chi<\pi \ \text{and} \ \operatorname{ctn}\chi=-\psi^{-1}\log\varrho .$$

Then, using (8), (11) and (9), it follows that

$$(13)\qquad \cos\theta_w^* = \frac{(p^*-w, v_1^*)}{|p^*-w|.|v_1^*|} = \cos\chi ,$$

and hence (as θ_w^* and χ are both in $[0,\pi]$)

$$(14)\qquad \theta_w^* = \chi .$$

We thus see that θ_w^* is indeed independent of w. Our condition (7^*) yields

$$(15)\qquad \chi \leqq \inf_{w\in U_o^*} \theta_{o,w}^* .$$

4 As σ is non-singular, W^* is like W an (n-1)-simplex. Hence U_o^*, being the boundary of the intersection of W^* with a plane that has at least one point (namely p^*) in common with W^*, is a convex polygon with at most n vertices, say $a_1,a_2,\dots,a_n$. It is then easy to see, that the infinum in the right-hand side of (15) is reached if w is in one of the vertices of U_o^*. Therefore, if we introduce 1)

$$\left.\begin{aligned} \alpha_j &= \measuredangle\, p^* a_j a_{j+1} \\ \alpha'_j &= \measuredangle\, p^* a_j a_{j-1} \end{aligned}\right\} \ (j=1,2,\dots,n;\ \text{indices taken mod } n),$$

then

$$(16)\qquad \inf_{w\in U_o^*} \theta_{o,w}^* = \min_j \alpha_j \ \text{or} \ \min_j \alpha'_j ,$$

----

1) For any three vectors a,b and c we use the following notation: $\measuredangle abc \overset{df}{=} \measuredangle (a-b,c-b)$.

depending on whether the sense of rotation in p^*+V^* given by p^*, w, v_1^* in that order is the same as the rotational sense given by $a_1, a_2, \ldots, a_n$ or that given by $a_n, a_{n-1}, \ldots, a_1$ respectively.

In 5 we shall give a proof of the fact that

$$\min_j \alpha_j \leqq \frac{n-2}{2n}\pi \quad ; \quad \min_j \alpha_j' \leqq \frac{n-2}{2n}\pi , \tag{17}$$

with equality in at least one of these two relations if and only if the polygon under consideration is regular and p^* its centre.

Combination of (15), (16) and (17) yields

$$\chi \leqq \frac{n-2}{2n}\pi ,$$

from which it follows that

$$\varrho \leqq e^{-\psi \operatorname{tg} \pi/n} . \tag{18}$$

Now ψ is of the form $\psi_0'+2k\pi$ with $0 \leqq \psi_0' < 2\pi$ and k an integer. As $\bar{\lambda}$ is also an eigenvalue of φ_1, we may conclude

$$\exists_k \quad \varrho \leqq e^{-(\psi_0'+2k\pi)\operatorname{tg} \pi/n} \text{ and } \varrho = e^{-(-\psi_0'-2k\pi)\operatorname{tg} \pi/n} . \tag{19}$$

If we put

$$\psi_0 = \min\{\psi_0', 2\pi-\psi_0'\} ,$$

than either $\psi_0'+2k\pi \geqq \psi_0$ or $-\psi_0'-2k\pi \geqq \psi_0$, hence

$$\varrho \leqq e^{-\psi_0 \operatorname{tg} \pi/n} . \tag{20}$$

If $\psi=0$, then $\psi_0=0$ also and as $\varrho \leqq 1$ always, (20) is trivially satisfied in this case. The restriction $\varrho \neq 0$ can now be removed also, as again (20) is satisfied in this case. Our final result is therefore: if λ is an eigenvalue of φ_1, ϱ its absolute value, ψ_0 the minimumvalue of the non-negative arguments of λ and $\bar{\lambda}$, then (20) must be satisfied. This result is equivalent to (6) in [1].

5 It remains to prove (17). To this purpose suppose, with the notation of section 4

(21) $$\forall_j \quad \alpha_j \geq \frac{n-2}{2n}\pi \ .$$

Then it is possible to find vectors[1] $p_1 \in \langle p^*, a_1\rangle$, $p_2 \in \langle p^*, a_2\rangle$, $\dots, p_n \in \langle p^*, a_n\rangle$ in such a way that
$\sphericalangle\, p^* a_n p_1 = \sphericalangle\, p^* p_1\, p_2 = \dots = \sphericalangle\, p^* p_{n-1} p_n = \frac{n-2}{2n}\pi$.
Then, as $p_n \in \langle p^* a_n\rangle$

$$\frac{|p_n - p^*|}{|a_n - p^*|} \leq 1,$$

with equality if and only if $\forall_j\ p_j = a_j$. Introducing $\alpha''_j = \sphericalangle\, p^* p_j p_{j-1}$ $(j=2,\dots,n)$ and $\alpha''_1 = \sphericalangle\, p^* p_1 a_n$, we find

$$1 \geq \frac{|p_n - p^*|}{|a_n - p^*|} = \frac{|p_n - p^*|}{|p_{n-1} - p^*|} \cdot \frac{|p_{n-1} - p^*|}{|p_{n-2} - p^*|} \cdots \frac{|p_1 - p^*|}{|a_n - p^*|} =$$

$$= \prod_{j=1}^{n} \frac{\sin \frac{n-2}{2n}\pi}{\sin \alpha''_j} \ .$$

Hence

(22) $$\prod_{j=1}^{n} \sin \alpha''_j \geq \sin^n \frac{n-2}{2n}\pi \ ,$$

with equality if and only if $\forall_j\ \alpha''_j = \alpha'_j$.

On the other hand, as the sum of the angles $(n\pi)$ of the n triangles $\langle a_n, p_1, p^*\rangle$, $\langle p_1, p_2, p^*\rangle$, $\dots$, $\langle p_{n-1}, p_n, p^*\rangle$ is equal to $\sum_{j=1}^{n} \alpha''_j + 2\pi + n.\frac{n-2}{2n}\pi$, we have

$$\sum_{j=1}^{n} \alpha''_j = n.\pi - 2\pi - \sum_{j=1}^{n} \frac{n-2}{2n}\pi = \frac{n-2}{2}\pi \ .$$

As $-\log \sin x$ is a convex function for $0 < x < \pi$, we have

----

1) We denote by $\langle b_1, b_2, \dots, b_k\rangle$ the closed convex hull of $\{b_1, \dots, b_k\}$: $\langle b_1, b_2, \dots, b_k\rangle = \{x \mid x = \alpha_1 b_1 + \alpha_2 b_2 + \dots + \alpha_k b_k;\ \forall_j\ \alpha_j \geq 0; \sum_{j=1}^{k} \alpha_j = 1\}$.

$$(23) \qquad \prod_{j=1}^{n} \sin \alpha_j'' \leqq \sin^n \frac{1}{n} \sum_{j=1}^{n} \alpha_j'' = \sin^n \frac{n-2}{2n}\pi ,$$

with equality if and only if $\forall_j \ \alpha_j'' = \frac{n-2}{2n}\pi$.
Combination of (22) and (23) yields:

$$(24) \qquad \forall_j \ \alpha_j' = \alpha_j'' = \frac{n-2}{2n}\pi .$$

We can now repeat this argument, starting from

$$(25) \qquad \forall_j \ \alpha_j' \geqq \frac{n-2}{2n}\pi ,$$

which follows from (24). We then arrive at the conclusion

$$(26) \qquad \forall_j \ \alpha_j = \frac{n-2}{2}\pi .$$

Hence, either of (21) or (25) entails (24) and (26). In other words: Either

$$\forall_j \ \alpha_j = \alpha_j' = \frac{n-2}{2n}\pi ,$$

in which case $\langle a_1, a_2, \ldots, a_n \rangle$ is regular with centre p^*, or

$$\min_j \ \alpha_j < \frac{n-2}{2n}\pi \ \text{ and } \min_j \ \alpha_j' < \frac{n-2}{2n}\pi ,$$

which proves (17).

References

[1] Runnenburg, J.Th., On K.L. Chung's problem of imbedding a time-discrete Markov-chain in a time-continuous one for finitely many states I, report S 299, Mathematisch Centrum, Amsterdam (1962).

[2] Dmitriev, N. and Dynkin E., On characteristic roots of stochastic matrices, Isvestija, ser.mat., 10 (1946) 167-194.

[3] Karpelewitsch, F.I. On the characteristic roots of a matrix with nonnegative elements, Isvestija, ser.mat., 15 (1951) 361-383.

[4] Fréchet, M., Méthode des fonctions arbitraires, Gauthier-Villars, Paris (1952).

[5] Kingman, J.F.C., The imbedding problem for finite Markov-chains, Z.f. Wahrscheinlichkeitstheorie und Verwandte Gebiete 1 (1962) 14-24.

[6] Elfving, G., Zur Theorie der Markoffschen Ketten, Acta Soc.Sci. Fennicae n.Ser. A 2, no 8 (1937).