

R. SYSKI (College Park, Md.)

PHASE-TYPE DISTRIBUTIONS AND PERTURBATION MODEL

1. INTRODUCTION

In recent papers [4]-[7], Neuts considered a finite Markov chain with one absorbing state, which is "re-started" whenever it enters the absorbing state. He termed the associated absolute distributions of time to absorption the phase-type (PH) distributions, and showed that their special properties make them useful in many applications (especially in queueing theory).

The purpose of this paper is to show that Neuts' construction can be interpreted in terms of the compensation method introduced earlier by Keilson (see [1] and [2]). The connection between these two developments is best explained with the help of the perturbation method, discussed by the present author in [8] and [9]. It turns out that Neuts' model is in fact a special form of the perturbation model.

It must be stressed that the comment above refers to the first stage of Neuts' work. Dealing with a specific perturbation model, he is able to obtain more detailed results concerning the class of PH distributions. Nevertheless, it is of interest to see the connection between the Neuts and Keilson approaches (applications of PH distributions discussed by Neuts are not considered here; see however [3] and [5]).

Section 2 describes the Neuts model (generalized to infinite chains with a single closed subset of states), and puts it in the framework of perturbation models. Section 3 describes the role of the compensation method in the study of this model. It is shown that the special structure of the Neuts model exhibits interesting features (Lemma 1 and Theorem 1) and that PH distributions are characterized by the special form of the replacement matrix whose elements are given by the initial distribution (Theorem 5). Theorems 2 and 4 summarize the role of the compensation method in the Neuts model, and the explicit expression for the compensation measure is given in Corollary 1. Section 4 contains proofs of these assertions.

Section 5 returns to the original Neuts model, and Section 6 presents illustrative examples.

2. THE MODEL

The aim of this section is to show that the model described by Neuts in [4]-[7] fits into the framework of perturbation models discussed in [8]. It is more convenient, however, to consider Markov chains with infinite state space. In the following, notation and terminology of [8] will be used.

2.1. Consider a continuous parameter Markov chain (M.C.) $\{X_t, t \geq 0\}$ with a discrete state space I , standard transition matrix $P(t) = (p_{ij}(t))$, infinitesimal generator $Q = (q_{ij})$ and resolvent $U^\alpha = (u_{ij}^\alpha)$. Denoting by $\mathbf{1}$ the column vector of 1's, we have $P(t) \cdot \mathbf{1} = \mathbf{1}$ and $Q \cdot \mathbf{1} = \mathbf{0}$. It is assumed that both the backward and the forward equations for $P(t)$ hold. See [8] for details.

Select a set $S \subset I$ of states, and write S^c for its complement. It is assumed that S^c is a *closed set* in the sense that no transitions from S^c to S are allowed (i.e., the M.C. cannot leave S^c , although transitions within S^c are permitted). On the other hand, the M.C. on S may stay in S or may leave S to get absorbed by S^c . It will be assumed that all states in S communicate. For convenience, it will be assumed that all states in S^c form a single communicating class, but this is not essential (all states in S^c being absorbing is a most notable exception).

Thus, the infinitesimal generator matrix Q and the transition matrix $P(t)$ take the partitioned forms

$$Q = \begin{array}{c} S^c \\ S \end{array} \begin{array}{cc} S^c & S \\ \left(\begin{array}{cc} Q_{11} & 0 \\ Q_{21} & Q_{22} \end{array} \right) \end{array} \quad \text{and} \quad P(t) = \begin{array}{c} S^c \\ S \end{array} \begin{array}{cc} S^c & S \\ \left(\begin{array}{cc} P_{11}(t) & 0 \\ P_{21}(t) & P_{22}(t) \end{array} \right) \end{array}.$$

In Neuts' original model, S is a finite set and S^c consists of a single absorbing state.

2.2. The following perturbation mechanism is now introduced in [4]-[7] resulting in a modified M.C. $\{X_t^*, t \geq 0\}$. Each time the original M.C. enters the closed set S^c , it is immediately "re-started" by returning to the set S . The returns are governed by the replacement matrix $R = (r_{ij})$ of the form

$$R = \begin{pmatrix} 0 & R_{12} \\ 0 & I \end{pmatrix}.$$

Here, r_{ij} is the probability that a state i is to be replaced by a state j . Clearly, $r_{ij} = 0$ for all $j \in S^c$ (because returns are only to states in S) and $r_{ij} = \delta_{ij}$ for $i \in S, j \in S$. Thus, proper returns are only from S^c to S described by the rectangular matrix $R_{12} = (r_{ij}), i \in S^c, j \in S$, of conditional proba-

bilities representing replacement distributions over S , for each $i \in S^c$, and

$$\sum_{j \in S} r_{ij} = 1, \quad i \in S^c.$$

(In Neuts' version, R_{12} reduces to a single row of one "re-starting" distribution.)

Of primary interest is the modified M.C. restricted to S . It comprises transitions of two kinds — those of the original M.C. governed by Q_{22} and the returns from S^c governed by $Q_{21}R_{12}$. Thus, the infinitesimal generator Q_{22}^* on S for the modified M.C. is of the form $Q_{22}^* = Q_{22} + Q_{21}R_{12}$. For convenience, one can consider the modified M.C. on the whole state space I by simply adjoining the isolated transitions on the closed set S^c . Thus, for the modified M.C. $\{X_t^*, t \geq 0\}$ we write

$$Q^* = \begin{pmatrix} Q_{11} & 0 \\ 0 & Q_{22}^* \end{pmatrix} \quad \text{and} \quad P^*(t) = \begin{pmatrix} P_{11}^*(t) & 0 \\ 0 & P_{22}^*(t) \end{pmatrix}.$$

It is clear that $P_{11}^*(t) = P_{11}(t)$.

It is the submatrix $P_{22}^*(t)$ which is of primary interest. Note that S is the closed set for the modified M.C. Furthermore, $Q^* \cdot 1 = 0$ and $P^*(t) \cdot 1 = 1$. Write also $U^{*a} = (u_{ij}^{*a})$ for the resolvent

$$U^{*a} = \begin{pmatrix} U_{11}^{*a} & 0 \\ 0 & U_{22}^{*a} \end{pmatrix}.$$

2.3. In conformity with Neuts' model, the following assumptions will be imposed:

- A1. The original M.C. on the closed set S^c is ergodic.
- A2. The modified M.C. on the closed set S is ergodic.
- A3. Q_{21} does not vanish identically.

Clearly, by construction, the original M.C. on S is necessarily transient.

Let $E = (e_{ij})$ and $E^* = (e_{ij}^*)$ be the limit matrices defined by

$$E = \lim_{t \rightarrow \infty} P(t) = \lim_{\alpha \rightarrow 0} \alpha U^a \quad \text{and} \quad E^* = \lim_{t \rightarrow \infty} P^*(t) = \lim_{\alpha \rightarrow 0} \alpha U^{*a}.$$

In the present case, under assumptions A1-A3, we have

$$E = \begin{pmatrix} E_{11} & 0 \\ E_{21} & 0 \end{pmatrix} \quad \text{and} \quad E^* = \begin{pmatrix} E_{11}^* & 0 \\ 0 & E_{22}^* \end{pmatrix}$$

with $E_{11} = E_{11}^*$, and $E \cdot 1 \leq 1$, $E^* \cdot 1 = 1$.

The rows of E_{11} are identical, giving the ergodic distribution $e = (e_j)$ on S^c . Similarly, the rows of E_{22}^* are identical, giving the ergodic distri-

bution $e^* = (e_j^*)$ on S for the modified chain:

$$\sum_{j \in S^c} e_j = 1, \quad \sum_{j \in S^c} e_{ij} \leq 1 \quad (i \in S), \quad \sum_{j \in S} e_j^* = 1,$$

where

$$\lim_{t \rightarrow \infty} p_{ij}(t) = e_{ij}, \quad i \in S, \quad j \in S^c.$$

2.4. In Neuts' presentation the principal role is played by the distribution of the first entrance time T to the closed set S^c for the original M.C., defined by

$$T = \inf\{t: t > 0, X_t \in S^c\}.$$

For any initial state $i \in I$, write for the conditional distribution of T :

$$D_i(t) = P(T \leq t \mid X_0 = i)$$

with $D_i(t) \equiv 1$ for $i \in S^c$ (by definition). As S^c is a closed set and the sample paths are right-continuous, it is clear that

$$p_{ij}(t) = P(X_t = j, T > t \mid X_0 = i), \quad i \in S, j \in S.$$

Indeed, this is the taboo probability of transition from i to j at t avoiding the taboo set S^c . Hence the complementary d.f. is given by

$$D_i^c(t) = \sum_{j \in S} p_{ij}(t), \quad i \in S,$$

and the conditional expectation $E(T, T < \infty \mid X_0 = i) = w_i$ assumed finite is clearly

$$w_i = \int_0^{\infty} [D_i^c(t) - D_i^c] dt, \quad i \in S.$$

For the probability of the first entrance to S^c we write

$$D_i = P(T < \infty \mid X_0 = i) \quad \text{and} \quad D_i^c = 1 - D_i.$$

Hence, for $i \in S$,

$$D_i^c = \lim_{t \rightarrow \infty} \sum_{j \in S} p_{ij}(t) \quad \text{and} \quad D_i = \sum_{j \in S^c} e_{ij} \leq 1.$$

Let π be the initial distribution on I for the original chain. The absolute distribution of T is then

$$F(t) = \sum_i \pi_i D_i(t) = \sum_{i \in S} \pi_i D_i(t) + \sum_{i \in S^c} \pi_i, \quad t \geq 0,$$

with

$$F(0) = \sum_{i \in S^c} \pi_i, \quad F(\infty) = F(0) + \sum_{i \in S} \pi_i D_i \leq 1$$

and the mean

$$W = \sum_{i \in S} \pi_i w_i = E(T, T < \infty).$$

The distribution F is the generalization of Neuts' PH distributions (see [4]-[7]).

The main problem of perturbation method is to express properties of the modified chain $P_{22}^*(t)$ on S in terms of properties of the original chain $P(t)$ on I . In particular, it is required to find the ergodic distribution e^* for the modified chain on S in terms of the characteristics of $P(t)$ on S .

In Neuts' analysis, the main problem is to express properties of the distribution F in terms of properties of the original chain $P(t)$ on I . In particular, it is required to relate the ergodic distribution e^* on S to the initial distribution π and the mean W of the PH distribution.

It is shown in the sequel that both problems are in fact two aspects of the same problem.

3. MAIN RESULTS

The purpose of this section is to show how the compensation method enters into the description of Neuts' model and relates to the PH distributions.

The general idea is the same as in [8] and it leads to the same second resolvent equation. However, in [8] it was assumed that all states in I communicate for the original chain; adaptation to the present situation where S^c is a closed set is immediate. The passage to the limit (as $a \rightarrow 0$) is now different. Nevertheless, the final result expressing the ergodic distribution for the modified chain as the Green potential of the compensation measure holds, and is the same as in [8]. Furthermore, it turns out that Neuts' method consists in fact in selecting the special form of the replacement matrix R .

For readers' convenience, this section contains only statements of the main results, and proofs and further comments are postponed to the next section.

3.1. Compensation method. Proceeding as in [8], define the time dependent compensation kernel $C(t) = (c_{ij}(t))$ by

$$C(t) = P^*(t)(Q^* - Q), \quad t \geq 0,$$

with $C(t) \cdot 1 = 0$. In the present case we have

$$C(t) = \begin{pmatrix} 0 & 0 \\ C_{21}(t) & C_{22}(t) \end{pmatrix},$$

where

$$C_{21}(t) = -P_{22}^*(t)Q_{21}, \quad C_{22}(t) = P_{22}^*(t)Q_{21}R_{12}.$$

Write

$$C^\alpha = \int_0^\infty e^{-at} C(t) dt, \quad \alpha > 0.$$

Then, under assumptions of Theorem 1.5 in [8], the second resolvent equation holds:

$$U^{*\alpha} = U^\alpha + C^\alpha U^\alpha.$$

This equation characterizes any perturbation model.

Define the limit compensation kernel $C = (c_{ij})$ by

$$C = \lim_{t \rightarrow \infty} C(t) = \lim_{\alpha \rightarrow 0} \alpha C^\alpha.$$

Under conditions of Theorem 1.8 in [8], C exists and

$$C = E^*(Q^* - Q) = -E^*Q$$

with $C \cdot 1 = 0$. In the present case we have

$$C = \begin{pmatrix} 0 & 0 \\ C_{21} & C_{22} \end{pmatrix}$$

with

$$C_{21} = -E_{22}^*Q_{21}, \quad C_{22} = E_{22}^*Q_{21}R_{12} = -E_{22}^*Q_{22}.$$

It is evident that rows of C do not depend on i , say $c_{ij} = c_j$ for every $i \in S, j \in I$. The row vector $c = (c_j)$ will be called the *compensation measure*.

The direct passage to the limit (as $\alpha \rightarrow 0$) in the second resolvent equation is not applicable, because by ergodicity the limit

$$\lim_{\alpha \rightarrow 0} U^\alpha = U$$

is infinite for all $j \in S^c$. It is therefore necessary to use the ergodic α -potential $Z^\alpha = (z_{ij}^\alpha)$ defined by

$$Z^\alpha = U^\alpha - \frac{E}{\alpha}, \quad \alpha > 0.$$

Arguments analogous to those in Theorem 1.14 in [8] (see also [9] and [10]) show the existence of the limit (called the *ergodic potential*)

$$Z = \lim_{\alpha \rightarrow 0} Z^\alpha$$

which in the present case is of the form

$$Z = \begin{pmatrix} Z_{11} & 0 \\ Z_{21} & U_{22} \end{pmatrix}$$

with $Z_{22} = U_{22}$, and $Z \cdot 1 = 0$.

As the M.C. is transient on S , U_{22} is finite. The matrix Z has only auxiliary character, and its properties are discussed in the next section.

The following lemma replaces Lemma 1.12 in [8] which is not applicable in the present situation.

LEMMA 1. *We have*

$$\lim_{\alpha \rightarrow 0} C^\alpha E = E^* E - E.$$

Using Lemma 1, the passage to the limit in the second resolvent equation yields

$$E^* = E^* E + CZ.$$

For the proof, see the next section. Hence, necessarily $E \cdot 1 = 1$.

This is the fundamental equation for the Neuts model, and it should be compared with equations (1.22) and (1.31) in [8]. It expresses the ergodic distribution on S for the modified chain in terms of the properties of the original chain. In fact, for transitions within S , the fundamental equation reduces to

$$E_{22}^* = C_{22} U_{22}.$$

Note that this expression has identical rows (independent of i). The matrix Z does not enter here, but it only refers to transitions from S to S^c through the submatrix Z_{21} . Furthermore, note that $U_{22} = (-Q_{22})^{-1}$.

Expressed in the component form, the discussion above can be summarized in the forthcoming theorem which is the main result of the compensation technique for the Neuts model. It should be compared with Theorems 1.10 and 1.14 in [8].

THEOREM 1. *For $P(t)$ transient on S and for $P^*(t)$ ergodic on S , the ergodic distribution $e^* = (e_j^*)$ on S (for the modified chain) is given explicitly by*

$$e_j^* = \sum_{k \in S} c_k u_{kj}, \quad j \in S,$$

where $c = (c_k)$ is the compensation measure (row vector of C), and $(u_{kj}) = U_{22}$ is the Green potential for transient $P_{22}(t)$ on S .

It is worthwhile to record properties of the compensation measure in the Neuts model.

THEOREM 2. *The compensation measure c has total mass zero, and its terms are positive on S and negative on S^c :*

$$\sum_j c_j = 0, \quad c_j \geq 0 \text{ for } j \in S, \quad c_j \leq 0 \text{ for } j \in S^c.$$

It should be remarked that assumption A2 of the ergodic e^* on S imposes constraints on the original M.C. Theorem 1 implies the following

THEOREM 3. *The time T is finite and*

$$D_i = \sum_{j \in S^c} e_{ij} = 1 \quad \text{for all } i \in S.$$

3.2. PH distributions. It is convenient to rewrite the definition of the distribution F in the form

$$F(t) = 1 - \sum_{i \in S} \pi_i D_i^c(t), \quad t \geq 0,$$

to stress that it depends essentially on the set S , with $F(0) \geq 0$, $F(\infty) = 1$, and the mean $W = E(T, T < \infty)$. By Theorem 3,

$$E(T | X_0 = i) = E(T, T < \infty | X_0 = i) = w_i$$

and

$$w_i = \sum_{j \in S} u_{ij} = - \sum_{j \in S^c} z_{ij}, \quad i \in S.$$

Note the interesting connection between the compensation measure and the conditional means:

THEOREM 4. *The conditional means w_i and the compensation measure $c_i \geq 0$ on S satisfy*

$$\sum_{i \in S} c_i w_i = 1.$$

As already noted, the first entrance distribution F depends essentially on the initial distribution π and on the perturbation mechanism through the replacement matrix R_{12} . The characterization of the Neuts PH distribution lies in the suitable choice of R_{12} .

Let $r = (r_j)$ be a distribution on S obtained by truncation of the initial distribution π on I to S ,

$$r_j = \frac{\pi_j}{\sum_{k \in S} \pi_k}, \quad j \in S.$$

Suppose now that the replacement matrix R_{12} has all rows identical and given by r : $r_{ij} = r_j$, $i \in S^c$, $j \in S$. The PH distributions are defined to be the first entrance distribution F with such a choice of the replacement

matrix. This in turn establishes the connection between ergodic distributions and compensation measures on one hand and between the initial distribution and PH distributions on the other hand.

The following theorem gives the important characterization of the Neuts model:

THEOREM 5. *Let π be the initial distribution on \mathbf{I} and let r be its truncation to S . Suppose that the replacement matrix R_{12} has all rows equal to r . Then the compensation measure c is given on S by*

$$c_j = \frac{\pi_j}{W}, \quad j \in S,$$

where W is the mean of the PH distribution F , and the ergodic distribution e^* on S is given by

$$e_j^* = \frac{1}{W} \sum_{i \in S} \pi_i u_{ij}, \quad j \in S.$$

4. PROOFS

Matrix notation is used for convenience, but all limit operations are taken pointwise. For any measure μ on \mathbf{I} and any function f on \mathbf{I} , write for their partition along S^c and S :

$$\mu = (\mu_1, \mu_2), \quad f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix},$$

and for their inner product:

$$\mu \cdot f = \mu_1 \cdot f_1 + \mu_2 \cdot f_2 = \sum_i \mu_i f_i.$$

In particular, $\pi = (\pi_1, \pi_2)$ and $c = (c_1, c_2)$ are the initial distribution and the compensation measure, respectively. In the same manner, $e_2^* = (e_j^*)$ is the ergodic distribution on S , and $w_2 = (w_i)$ is the mean vector on S . Define also the function $D(\cdot)$ by

$$D(t) = \begin{pmatrix} D_1(t) \\ D_2(t) \end{pmatrix} = (D_i(t)) \quad \text{with } D_1(t) \equiv 1;$$

similarly for $D = (D_i)$.

4.1. Proof of Lemma 1. Write the forward equation for $P^*(t)$ in the transformed form, as in [8]:

$$\alpha U^{*a} - I = U^{*a} Q + C^a.$$

Post-multiplying by E and noting that always $QE = 0$, we have

$$\alpha U^{*\alpha} E - E = C^\alpha E.$$

The passage to the limit as $\alpha \rightarrow 0$ (justified by the Helly-Bray theorem) yields the required result. Note that in the partitioned form all submatrices vanish, except for $i \in S, j \in S^c$, so Lemma 1 reduces to

$$C_{21}^\alpha E_{11} + C_{22}^\alpha E_{21} \rightarrow E_{22}^* E_{21} - E_{21},$$

where C_{21}^α and C_{22}^α both diverge as $\alpha \rightarrow 0$.

4.2. Proof of the fundamental equation. Using the expression for Z^α , rewrite the second resolvent equation in the form

$$\alpha U^{*\alpha} = \alpha U^\alpha + C^\alpha \alpha U^\alpha = \alpha U^\alpha + \alpha C^\alpha Z^\alpha + C^\alpha E.$$

As in Theorem 1.14 in [8] (see also [9]),

$$\lim_{\alpha \rightarrow 0} \alpha C^\alpha Z^\alpha = CZ.$$

By Lemma 1, the passage to the limit yields

$$E^* = E^* E + CZ.$$

4.3. Ergodic potential. In the present case

$$Z^\alpha = \begin{pmatrix} Z_{11}^\alpha & 0 \\ Z_{21}^\alpha & Z_{22}^\alpha \end{pmatrix},$$

where

$$Z_{11}^\alpha = U_{11}^\alpha - \frac{E_{11}}{\alpha}, \quad Z_{21}^\alpha = U_{21}^\alpha - \frac{E_{21}}{\alpha}, \quad Z_{22}^\alpha = U_{22}^\alpha.$$

Indeed, $E_{22} = 0$ and Z_{11}^α is precisely the ergodic α -potential discussed in [9]. On the other hand, Z_{21}^α is a new quantity which appears in the present situation.

By definition, $Z_{11}^\alpha \cdot 1 = 0$, but $E_{21} \cdot 1 \leq 1$ implies

$$Z_{21}^\alpha \cdot 1 + Z_{22}^\alpha \cdot 1 = \frac{1}{\alpha} - \frac{1}{\alpha} E_{21} \cdot 1 \geq 0,$$

so $Z^\alpha \cdot 1 \geq 0$.

4.4. Proof of Theorem 1. Note that

$$CZ = \begin{pmatrix} 0 & 0 \\ C_{21} Z_{11} + C_{22} Z_{21} & C_{22} U_{22} \end{pmatrix}, \quad E^* E = \begin{pmatrix} E_{11} & 0 \\ E_{22}^* E_{21} & 0 \end{pmatrix}.$$

Hence the fundamental equation $E^* = E^* E + CZ$ yields

$$E_{22}^* E_{21} + C_{21} Z_{11} + C_{22} Z_{21} = 0, \quad E_{22}^* = C_{22} U_{22}.$$

Observe that both these equations have rows independent of i . When written in the component form, the formulae in Theorem 1 are obtained.

4.5. Proof of Theorem 2. By the definition of C , clearly $c \cdot 1 = 0$. The positivity on S follows from

$$0 = e_2^* Q_{22}^* = e_2^* Q_{22} + e_2^* Q_{21} R_{12} \quad \text{and} \quad Q_{21} R_{12} \geq 0,$$

so $c_2 = -e_2^* Q_{22} \geq 0$. Similarly, $C_{11} = -E_{22}^* Q_{21} \leq 0$, so $c_1 \leq 0$.

4.6. Proof of Theorem 3. By assumption, $E^* \cdot 1 = 1$ and $Z \cdot 1 = 0$. Hence, from the fundamental equation we get

$$E^* \cdot 1 = E^* E \cdot 1 + CZ \cdot 1 = E^*(E \cdot 1).$$

As $E \cdot 1 \leq 1$, we obtain necessarily $E \cdot 1 = 1$ and, in particular, $E_{21} \cdot 1 = 1$. But $E_{21} \cdot 1 = D_2$, so $D_2 = 1$. Consequently, $F(\infty) = 1$.

4.7. As concerns the first entrance distributions in Section 2.4 we can write

$$D_2(t) = P_{21}(t) \cdot 1, \quad D_2^c(t) = P_{22}(t) \cdot 1, \quad F^c(t) = \pi_2 \cdot D_2^c(t).$$

Furthermore, the mean vector on S is $w_2 = U_{22} \cdot 1$ and the mean W of the distribution F is clearly the inner product $W = \pi_2 \cdot w_2$. The definition of F can be written as

$$F(t) = 1 - \pi_2 P_{22}(t) \cdot 1$$

with

$$F(0) = 1 - (\pi_2 \cdot 1), \quad F(\infty) - F(0) = \pi_2 \cdot D_2.$$

4.8. The vector (d_i^a) of the Laplace-Stieltjes transforms of $D_i(t)$ is defined by

$$d_i^a = \int_0^\infty e^{-at} dD_i(t), \quad a > 0, \quad i \in S.$$

It is easy to see that

$$d_2^a = a U_{21}^a \cdot 1 = 1 - a U_{22}^a \cdot 1.$$

Hence

$$Z_{21}^a \cdot 1 = \frac{1}{a} (d_2^a - D_2)$$

because $E_{21} \cdot 1 = D_2$. Hence, passing to the limit ($a \rightarrow 0$) we get $Z_{21} \cdot 1 = -w_2$. But $Z_{22} \cdot 1 = U_{22} \cdot 1 = w_2$, so $Z_{21} \cdot 1 + Z_{22} \cdot 1 = 0$. Obviously, $Z_{11} \cdot 1 = 0$.

4.9. Proof of Theorem 4. By Theorems 1 and 2,

$$1 = e_2^* \cdot 1 = c_2 U_{22} \cdot 1 = c_2 \cdot U_{22} \cdot 1 = c_2 \cdot w_2.$$

(Note that selecting $\pi_2 = c_2$ would give $W = 1$.)

4.10. Proof of Theorem 5. By Theorem 1, $c_2 = e_2^* Q_{21} R_{12}$, but now R_{12} is a matrix of identical rows r_2 , so

$$c_2 = (e_2^* Q_{21} \cdot 1) r_2.$$

On the other hand,

$$R_{12} w_2 = (r_2 \cdot w_2) \cdot 1,$$

so by Theorem 4 we have

$$1 = c_2 \cdot w_2 = e_2^* Q_{21} R_{12} w_2 = (e_2^* Q_{21} \cdot 1) (r_2 \cdot w_2)$$

or

$$c_2 = \frac{r_2}{r_2 \cdot w_2} = \frac{\pi_2}{\pi_2 \cdot w_2}$$

because $r_2 = \pi_2 / (\pi_2 \cdot 1)$. The expression for e_2^* follows then from Theorem 1.

4.11. Remarks. 1. Consider now the variant of the Neuts model in which all states in S^c are assumed to be absorbing (thus assumption A1 is dropped).

Hence $Q_{11} = 0$ and, clearly, $P_{11}(t) = E_{11} = P_{11}^*(t) = E_{11}^* = I$. The expressions for C^a and C do not change, however. On the other hand, $Z_{11}^a = Z_{11} = 0$. Lemma 1 does not change, and the fundamental equation holds. Theorem 1 yields as before $E_{22}^* = C_{22} U_{22}$, but the second equation reduces to $E_{22}^* E_{21} + C_{22} Z_{21} = 0$.

2. Define the first entrance distribution by

$$D_{ij}(t) = P(X_T = j, T \leq t \mid X_0 = i), \quad i \in S, j \in S^c.$$

Then

$$D_i(t) = \sum_{j \in S^c} D_{ij}(t).$$

Observe, however, that $D_{ij}(t) \neq p_{ij}(t)$ for $i \in S, j \in S^c$, because $p_{ij}(t)$ involves both the first entry to S^c from S and the subsequent transitions within S^c . Nevertheless

$$\sum_{j \in S^c} D_{ij}(t) = \sum_{j \in S^c} p_{ij}(t),$$

since S^c is a closed set ($Q_{11} \cdot 1 = 0$).

If it were assumed that all states of S^c are absorbing, then necessarily $D_{ij}(t) = p_{ij}(t)$ for all $j \in S^c$.

5. NEUTS' MODEL

It is of interest to specialize the results from Sections 2-4 to the model treated by Neuts in [4]-[7].

As already noted, Neuts considered a finite chain with $m+1$ states, where $S = (1, \dots, m)$, and S^c consisting of a single state, necessarily absorbing. Neuts denoted the absorbing state by $m+1$, but it is more convenient here to denote it by 0, so $S^c = (0)$.

Consequently, the replacement matrix R_{12} reduces to a single row of the re-starting distribution r :

$$r_j = \frac{\pi_j}{1-r_0}, \quad j = 1, \dots, m,$$

where π is the initial distribution. Thus, the situation of Theorem 5 prevails. In Neuts' notation: $\pi = (\alpha, \alpha_{m+1})$ with $\alpha_{m+1} = \pi_0$ taken as 0 for most of the discussion.

As the chain is finite, we have the representation

$$P_{22}(t) = \exp(Q_{22}t) \quad \text{with} \quad U_{22} = (-Q_{22})^{-1}.$$

(Q_{22} is T in Neuts' notation.)

On the other hand, $P_{21}(t) \rightarrow E_{21}$ are column vectors such that

$$p_{i0}(t) \rightarrow e_{i0} = 1, \quad i = 1, \dots, m,$$

so $D_i = 1$ identically. Indeed, the absorption into a state 0 from any initial state $i \in S$ is certain in a finite chain.

The matrix Q_{22}^* has elements $q_{ij}^* = q_{ij} + q_{i0}r_j$ ($i, j = 1, \dots, m$) and the ergodic potential is found to be

$$z_{00} = 0, \quad z_{i0} = - \sum_{j=1}^m u_{ij}, \quad i = 1, \dots, m, \quad z_{ij} = u_{ij}, \quad i, j = 1, \dots, m,$$

with

$$\sum_{j=1}^m z_{ij} = 0 \quad \text{for each } i.$$

The definition of the PH distribution used by Neuts takes the form

$$F(t) = 1 - \pi_2 \exp(Q_{22}t) \cdot \mathbf{1}$$

(where π_2 is the restriction of π to S) with $F(0) = \pi_0$, $F(\infty) = 1$ and the mean

$$W = \pi_2 \cdot U_{22} \cdot \mathbf{1} = \pi_2 (-Q_{22})^{-1} \cdot \mathbf{1}.$$

Theorems 1 and 5 combined yield immediately

COROLLARY 1. *The ergodic distribution is given by*

$$e_j^* = \sum_{k=1}^m c_k u_{kj} = \frac{1}{W} \sum_{k=1}^m \pi_k u_{kj}, \quad j = 1, \dots, m,$$

where the compensation measure is of the form

$$c_j = \frac{\pi_j}{W}, \quad j = 1, \dots, m, \quad c_0 = -\frac{1 - \pi_0}{W}.$$

6. EXAMPLES

The examples in this section have been selected for their simplicity in order to produce the explicit expressions for the compensation measure, ergodic distribution and PH distributions.

Example 1 refers to a finite chain, and all others to infinite chains. All examples except Example 4 have S^0 consisting of a single absorbing state. In Example 4 the closed set S^0 contains two states.

Except perhaps for Example 4, the other examples are standard and well known. Example 1 was used by Neuts. Examples 2 and 4 show interesting similarities, confirming that the set S^0 has not much influence on Neuts' formulation. Example 3 is the interesting modification of the classical case.

Example 1. *Erlang distribution.*

Let $S = (1, \dots, m)$ and let 0 be the absorbing state. Suppose that transitions on S follow the (truncated) Poisson process with a parameter λ , but the last state m leads to the absorbing state 0. The matrix Q is now of the form

$$Q = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & \dots & m-1 & m \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \dots \\ m-1 \\ m \end{matrix} & \left(\begin{array}{c|cccccc} \hline 0 & 1 & 2 & \dots & m-1 & m \\ \hline 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & -\lambda & \lambda & \dots & 0 & 0 \\ 0 & 0 & -\lambda & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -\lambda & \lambda \\ \lambda & 0 & 0 & \dots & 0 & -\lambda \end{array} \right) \end{matrix}.$$

It is easy to see that solutions for the transition probabilities and their transforms are

$$p_{ij}(t) = \frac{(\lambda t)^{j-i}}{(j-i)!} e^{-\lambda t}, \quad u_{ij}^\alpha = \left(\frac{\lambda}{\lambda + \alpha} \right)^{j-i} \frac{1}{\lambda + \alpha}$$

for $i = 1, \dots, m, j = 1, \dots, m$ and $j \geq i$ (and 0 otherwise), and

$$p_{i0}(t) = 1 - e^{-\lambda t} \sum_{k=0}^{m-i} \frac{(\lambda t)^k}{k!} = \frac{1}{(m-i)!} \int_0^{\lambda t} x^{m-i} e^{-x} dx,$$

$$u_{i0}^\alpha = \frac{1}{\alpha} - \frac{1}{\lambda} \sum_{k=0}^{m-i} \left(\frac{\lambda}{\lambda + \alpha} \right)^k$$

for $i = 1, \dots, m$.

Clearly, $D_i(t) = p_{i0}(t)$ is the Erlang distribution with $D_i \equiv 1$, and the mean $w_i = (m - i + 1)/\lambda$. Furthermore, $z_{i0} = -w_i$, $u_{ij} = 1/\lambda$ for $j \geq i$.

Take the initial distribution π with $\pi_0 = 0$, and let γ be the mean of π . Then the PH distribution

$$F(t) = \sum_{i=1}^m \pi_i p_{i0}(t)$$

has a density which is a linear combination of Erlang densities, and the mean is

$$W = \sum_{i=1}^m \pi_i w_i = (m - \gamma + 1)/\lambda.$$

Taking π as the replacement distribution (as in Section 5), we obtain the matrix Q_{22}^* in the form

$$Q_{22}^* = \begin{matrix} & \begin{matrix} 1 & 2 & \dots & m-1 & m \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ \dots \\ m-1 \\ m \end{matrix} & \begin{pmatrix} -\lambda & \lambda & \dots & 0 & 0 \\ 0 & -\lambda & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \lambda\pi_1 & \lambda\pi_2 & \dots & \lambda\pi_{m-1} & -\lambda(1-\pi_m) \end{pmatrix} \end{matrix}.$$

The ergodic distribution e^* on S can be found directly from $e^*Q_{22}^* = 0$. However, it is easy to see that

$$e_j = \frac{\pi_j}{W}, \quad j = 1, \dots, m,$$

and therefore

$$e_j^* = \frac{\pi_1 + \dots + \pi_j}{m+1-\gamma}, \quad j = 1, \dots, m.$$

Note that if returns are always made to a fixed state, say k , then $F(t) = p_{k0}(t)$ and e^* is the uniform distribution $1/(m+1-k)$.

Example 2. *Death process.*

Let $S = (1, 2, \dots)$ be the set of positive integers, and let 0 be the absorbing state. Suppose that transitions on S follow the death process with $q_{i,i-1} = \mu$, so absorption occurs only through the state 1. The matrix Q is now of the form

$$Q = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & \dots \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ \dots \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ \mu & -\mu & 0 & 0 & \dots \\ 0 & \mu & -\mu & 0 & \dots \\ 0 & 0 & \mu & -\mu & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \end{matrix}.$$

It is easy to see that the transition probabilities are defined by

$$p_{ij}(t) = \frac{(\mu t)^{i-j}}{(i-j)!} e^{-\mu t} \quad \text{for } i \geq 1, 0 < j \leq i,$$

with

$$u_{ij} = \frac{1}{\mu} \quad \text{for } i \geq j$$

and vanish for $i < j$.

The absorption probabilities for $i \geq 1$ are

$$p_{i0}(t) = 1 - e^{-\mu t} \sum_{k=0}^{i-1} \frac{(\mu t)^k}{k!} = \frac{1}{(i-1)!} \int_0^{\mu t} x^{i-1} e^{-x} dx.$$

Clearly, $D_i(t) = p_{i0}(t)$, and

$$D_i = \lim_{t \rightarrow \infty} p_{i0}(t) = e_{i0} = 1$$

and the mean is $w_i = i/\mu$.

Furthermore, the ergodic potential is

$$z_{00} = 0, \quad z_{i0} = -i/\mu, \quad u_{ij} = 1/\mu \quad \text{for } i \geq j \ (i \geq 1).$$

The replacement matrix R_{12} reduces to the single row vector $r = (r_j)$. Hence

$$Q_{22}^* = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & \dots \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ \dots \end{matrix} & \begin{pmatrix} -\mu(1-r_1) & \mu r_2 & \mu r_3 & \dots \\ \mu & -\mu & 0 & \dots \\ 0 & \mu & -\mu & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \end{matrix}.$$

The solution of $e^*Q_{22}^* = 0$ is easily found to be

$$e_j^* = \frac{1}{\gamma} \sum_{k=j}^{\infty} r_k, \quad j = 1, 2, \dots,$$

where γ is the mean of distribution r .

On the other hand, Theorem 1 gives

$$e_j^* = \frac{1}{\mu} \sum_{k=j}^{\infty} c_k, \quad j = 1, 2, \dots,$$

so necessarily

$$c_j = \frac{\mu}{\gamma} r_j, \quad j = 1, 2, \dots, \quad c_0 = -e_1^* q_{10} = -e_1^* \mu = -\frac{\mu}{\gamma},$$

and, consequently,

$$\sum_{j=0}^{\infty} c_j = 0.$$

The same result may be obtained by Neuts' argument (as in Section 5). Take r for the initial distribution concentrated on S ($\pi_j = r_j$ for $j \geq 1$). Then the PH distribution

$$F(t) = \sum_{i=1}^{\infty} \pi_i p_{i0}(t)$$

has a density which is a weighted sum of Erlang densities, and the mean is

$$W = \sum_{i=1}^{\infty} \pi_i w_i = \frac{\gamma}{\mu}.$$

Hence, by Theorem 5,

$$c_j = \frac{\pi_j}{W}, \quad j = 1, 2, \dots, \quad c_0 = -\frac{1}{W},$$

and

$$e_j^* = \frac{1}{\mu W} \sum_{i=j}^{\infty} \pi_i, \quad j = 1, 2, \dots,$$

as before.

Example 3. Busy period for $M/M/1$.

Consider the usual $M/M/1$ queue, and let 0 be the absorbing state. Here $S = (1, 2, \dots)$, $S^c = (0)$, and λ and μ are the usual birth and death coefficients. For $i, j = 1, 2, \dots$, $p_{ij}(t)$ are the taboo probabilities with 0

as the taboo state. On the other hand, $p_{i0}(t) = D_i(t)$, $i = 1, 2, \dots$, is the distribution of the busy period (initiated by the state i). As is well known:

$$p_{i0}(t) \rightarrow D_i = \begin{cases} \left(\frac{\mu}{\lambda}\right)^i & \text{if } \lambda \geq \mu, \\ 1 & \text{if } \lambda < \mu. \end{cases}$$

The replacement matrix R_{12} reduces to the single row vector $r = (r_j)$. Hence

$$Q_{22}^* = \begin{matrix} & \begin{matrix} 1 & 2 & 3 & 4 & \dots \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ \dots \end{matrix} & \begin{pmatrix} -\lambda - \mu(1-r_1) & \lambda + \mu r_2 & \mu r_3 & \mu r_4 & \dots \\ \mu & -(\lambda + \mu) & \lambda & 0 & \dots \\ 0 & \mu & -(\lambda + \mu) & \lambda & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \end{matrix}$$

In order to solve $e^* Q_{22}^* = 0$ for the ergodic distribution, introduce the generating functions

$$\psi(z) = \sum_{j=1}^{\infty} e_j^* z^j \quad \text{and} \quad \beta(z) = \sum_{j=1}^{\infty} r_j z^j.$$

Then one finds by routine calculations that the ergodic distribution exists provided

$$\rho = \frac{\lambda}{\mu} < 1$$

and that

$$\psi(z) = \frac{z}{\gamma} \frac{1-\rho}{1-\rho z} \frac{1-\beta(z)}{1-z},$$

where γ is the mean of distribution r .

Note that $\psi(z)$ involves the geometric distribution and the tail of distribution r . Hence, by inversion,

$$e_1^* = \frac{1-\rho}{\gamma}$$

and

$$e_j^* = \frac{1}{\gamma} \left[1 - \rho^j - \sum_{n=1}^{j-1} r_n (1 - \rho^{j-n}) \right], \quad j = 2, 3, \dots$$

It is of interest to note that in the special case where the M.C. is always re-started at $i = 1$ ($r_1 = 1$) the ergodic distribution is geometric: $e_j^* = (1-\rho) \rho^{j-1}$.

Although the explicit expressions for $p_{i0}(t)$ are cumbersome, it is known that its mean is

$$w_i = \frac{i}{\mu(1-\rho)}, \quad i = 1, 2, \dots$$

(for $\rho < 1$).

Take now the initial distribution to coincide with r (as in Section 5). Then the PH distribution

$$F(t) = \sum_{i=1}^{\infty} r_i p_{i0}(t)$$

has the mean

$$W = \sum_{i=1}^{\infty} r_i w_i = \frac{\gamma}{\mu(1-\rho)}.$$

Hence, by Theorem 5, the compensation measure is

$$c_j = \frac{r_j}{W}, \quad j = 1, 2, \dots, \quad \text{with } c_0 = -\frac{1}{W}.$$

By Theorem 5,

$$e_j^* = \frac{1}{W} \sum_{i=1}^{\infty} r_i u_{ij}, \quad j = 1, 2, \dots$$

It is now necessary to find the quantities u_{ij} . They can be obtained from the relation $U_{22}Q_{22} = -I$.

Rather cumbersome calculations give, for $i, j = 1, 2, \dots$,

$$u_{ij} = \begin{cases} \frac{1-\rho^j}{\mu(1-\rho)}, & j = 1, 2, \dots, i, \\ \frac{(1-\rho^i)\rho^{j-i}}{\mu(1-\rho)}, & j = i+1, \dots, \end{cases}$$

for $\rho < 1$.

It can be verified that

$$w_i = \sum_{j=1}^{\infty} u_{ij} = -z_{i0}, \quad i = 1, 2, \dots$$

The substitution of u_{ij} into the formula above for e_j^* results in the same expressions found directly earlier.

Example 4. Two-state closed set.

This example is the extension of Example 2 to the case where the closed set S^c consists of two states.

In connection with Remark 2 in Section 4.11 it is of interest to note that $D_{i0}(t) \equiv 0$, but $D_{i1}(t) = D_i(t)$, confirming that

$$D_{ij}(t) \leq p_{ij}(t), \quad j = 0, 1.$$

Suppose now that the initial distribution π is concentrated on S , and let γ be its mean and Π its generating function,

$$\Pi(z) = \sum_{i=2}^{\infty} \pi_i z^i.$$

Then the PH distribution

$$F(t) = \sum_{i=2}^{\infty} \pi_i D_i(t)$$

has the Laplace-Stieltjes transform

$$\int_0^{\infty} e^{-at} dF(t) = \sum_{i=2}^{\infty} \pi_i d_i^a = \frac{\mu + a}{\mu} \Pi\left(\frac{\mu}{\mu + a}\right)$$

with the mean

$$W = \sum_{i=2}^{\infty} \pi_i w_i = \frac{\gamma - 1}{\mu}.$$

In the present case the replacement matrix R_{12} has two rows. In agreement with Theorem 5, suppose that these rows are the same and equal to π . Then, by Theorem 5, the compensation measure is

$$c_j = \frac{\pi_j}{W}, \quad j = 2, 3, \dots,$$

and, consequently, the ergodic distribution is

$$e_j^* = \sum_{k=j}^{\infty} c_k u_{kj} = \frac{1}{\mu W} \sum_{k=j}^{\infty} \pi_k, \quad j = 2, 3, \dots$$

Note that it is of the same form as in Example 2, the difference being that γ is replaced by $\gamma - 1$.

Finally, note that

$$c_0 = 0, \quad c_1 = -\frac{1 - \pi_1}{W},$$

so

$$\sum_{j=0}^{\infty} c_j = 0.$$

Acknowledgment. The author wishes to express his thanks to Julian Keilson and Marcel Neuts for their helpful comments and for preprints of their papers.

Added in proof. Perturbation technique as a random time transformation has been described in the paper: P. D. Feigin and E. Rubinstein, *Equivalent descriptions of perturbed Markov processes*, Stochastic Processes Appl. 9 (1979), p. 261-272.

References

- [1] J. Keilson, *The role of Green's functions in congestion theory*, p. 43-71 in: W. L. Smith and W. E. Wilkinson (Eds.), *Congestion theory*, Univ. of North Carolina Press, Chapel Hill, N. C., 1965.
- [2] — *Markov chain models — rarity and exponentiality*, Applied Mathematical Sciences No. 28, Springer-Verlag, New York - Berlin 1979.
- [3] — and W. R. Nunn, *Laguerre transformation as a tool for the numerical solution of integral equations of convolution type*, Appl. Math. Comput. 8 (1981), p. 137-174.
- [4] M. F. Neuts, *Probability distribution of phase type*, p. 173-206 in: Liber Amicorum Professor Emeritus H. Florin, Dept. of Mathematics, University of Louvain, Belgium, 1975.
- [5] — *Computational uses of the method of phases in the theory of queues*, Comput. Math. Appl. 1 (1975), p. 151-166.
- [6] — *Markov chains with applications in queueing theory, which have a matrix-geometric invariant probability vector*, Advances in Appl. Probability 10 (1978), p. 185-212.
- [7] — *Renewal processes of phase type*, Naval Res. Logist. Quart. 25 (1978), p. 445-454.
- [8] R. Syski, *Perturbation models*, Stochastic Processes Appl. 5 (1977), p. 93-129.
- [9] — *Ergodic potential*, ibidem 7 (1978), p. 311-336.
- [10] — *Energy for Markov chains*, Advances in Appl. Probability 11 (1979), p. 542-575.

DEPT. OF MATHEMATICS
UNIVERSITY OF MARYLAND
COLLEGE PARK, MD. 20742, U.S.A.

Received on 25. 8. 1979

R. SYSKI (College Park, Md.)

ROZKŁADY TYPU FAZOWEGO I MODEL PERTURBACJI

STRESZCZENIE

W ostatnich latach Neuts wprowadził tzw. *rozkłady typu fazowego*, otrzymane jako rozkłady czasu do absorpcji w skokowym procesie Markowa (z czasem ciągłym, skończoną liczbą stanów i jednym stanem absorpcyjnym) przy dodatkowym założen-

niu, że proces zostaje zaczęty od nowa, gdy ulegnie absorpcji. Wcześniej już Keilson wprowadził metodę kompensacji celem wyrażenia charakterystyk procesów Markowa przy użyciu funkcji pomocniczych.

Syski zastosował metodę kompensacji do analizy łańcuchów Markowa (czas ciągle, przeliczalna liczba stanów i wszystkie stany istotne), w których występują zakłócenia. Mechanizm perturbacji polega na tym, że gdy łańcuch znajduje się w grupie stanów „dozwolonych”, wtedy przebiega bez zakłóceń, lecz gdy znajdzie się w grupie stanów „zabronionych”, zostaje natychmiast przeniesiony (według rozkładu rozmieszczenia) do grupy stanów dozwolonych. Otrzymuje się w ten sposób dwa łańcuchy i problem polega na powiązaniu ich własności.

W pracy wykazuje się, iż konstrukcja Neutsa jest w zasadzie specjalnym przypadkiem metody perturbacji, scharakteryzowanej przez pewną postać rozkładu rozmieszczenia określonego przez rozkład początkowy. Powiązanie to jest wyjaśnione za pomocą metody kompensacji. Rozdział 2 zawiera krótkie, lecz wyczerpujące wprowadzenie do metody perturbacji i do rozkładów typu fazowego. Główne rezultaty zawarte są w rozdziale 3, który omawia rolę metody kompensacji w analizie tych rozkładów oraz podaje ich własności. Rozkłady typu fazowego omawiane tutaj są uogólnieniem rozkładów Neutsa do przeliczalnej liczby stanów i do dowolnego odizolowanego zbioru stanów zabronionych. Rozdział 4 podaje dowody twierdzeń. Rozdział 5 zacieśnia rozważania do pierwotnego modelu Neutsa, a rozdział 6 zawiera przykłady.
